

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACCTC
AGTGAAGTGAAGGCCAGGCCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCTCGCCGCTGTG
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTCCGTACCTGACCTCCATCCCAG
ATGGCATCCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC
TAAAAAATGAGCTATAACAAAGTCCAAATCATTCCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCCAAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCCTTTGCTATGGTACCATCTGGAGCTTTCTCTATGTACAA
AGCCAACCATTTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACCTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCTCAGGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGTTGGGATTCCACG
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCAACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAACCTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCCAGATGAAGAATTCATGGTTC
TGAAAATAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGACACCCCTGTAACAGGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTAGAAACAA
CAATACAGGTAACCTTCCCCTTGTCCAGGCACCTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAATA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAACTGGGAAAGATTGAGTGGAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACCTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAAGTCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCTTGGCTTTTTCAGTCATTCCCCACGACAAGATGATGGTGGAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCTCCTCAGC
AGTGGGGGCGCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAACCCAACCTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAACATAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACCTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

Figure 2

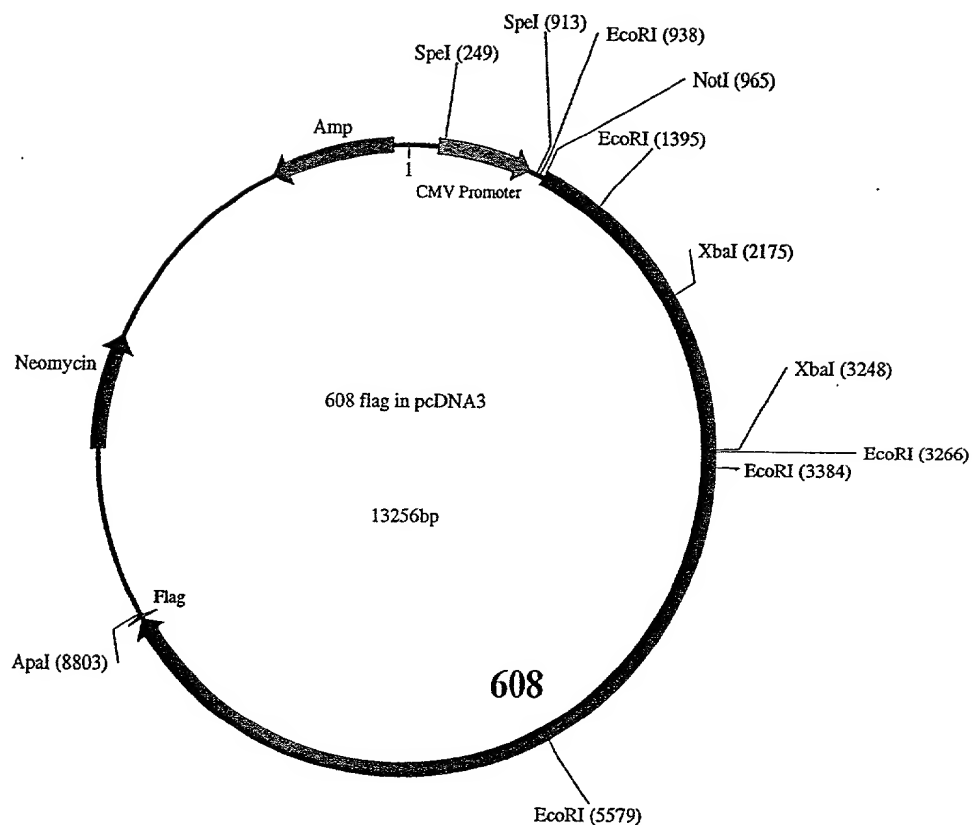


Figure 3

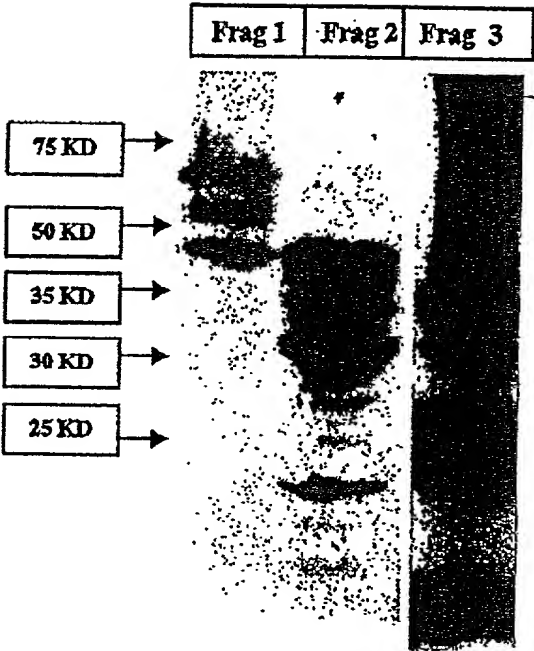
MQVRGREVSGLLISLTA VCLVVTPGSRACPRRCACYVPTEVHCTFRYLT SIPDGIPANVE
RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSLTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRK
NKHRLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRHRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGTLVATAALS
VPSSSHSALPKTNVGVIAEESTTVVKKPLLLFKDKQNVDIEIITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHLSPGTIQTGKDSVETPLPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
TMAATQNKGTENVSGARSLSAGKKQPFTNSSPVLPTISKRSNTLNFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQESTTVKRASDTP
PPLLSSGAPPVPTSPPPFTKGVVTD SKVTSAFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSPWPEHKYQLKSYSETIEKGKRP
VSMSPHLSLPEASTHASHWNTQKHAEKSVFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANSDVFLPCEAVGDPLPIHWTRVSSGXEISQGTQKSRFHVLPNGTSLIQRV
SIQDRGQYLCSAFNPLGVDHFHVLSLVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGTLLIYNLSLYDRGFYKCVASNPSGQDSSL
VKIQVITAPPVIEQKRQAIVGVLGGSLLKPCTAKGTPQPSVHWVLYDGTTELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVVILTVEEGETIPRIETASQKWTEVN
LGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRMGSRHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTL SIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLLYKATRKNKSGKYRCAARNKVGYYEKLIL
LEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVAYPPIINYLPNNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4



9/90
Figure 5

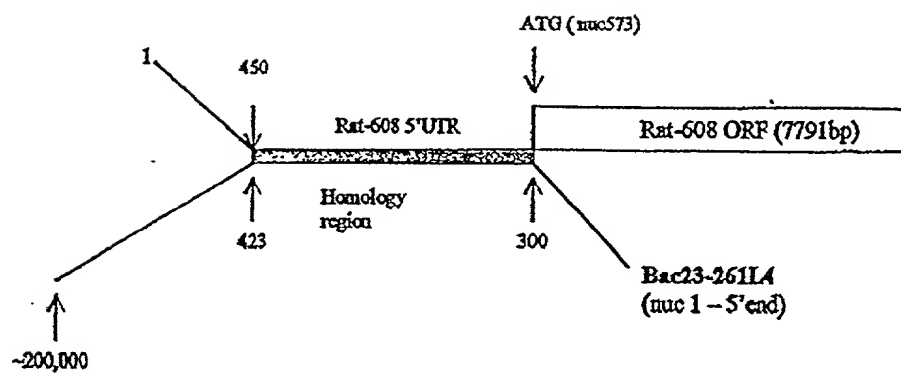


Figure 6

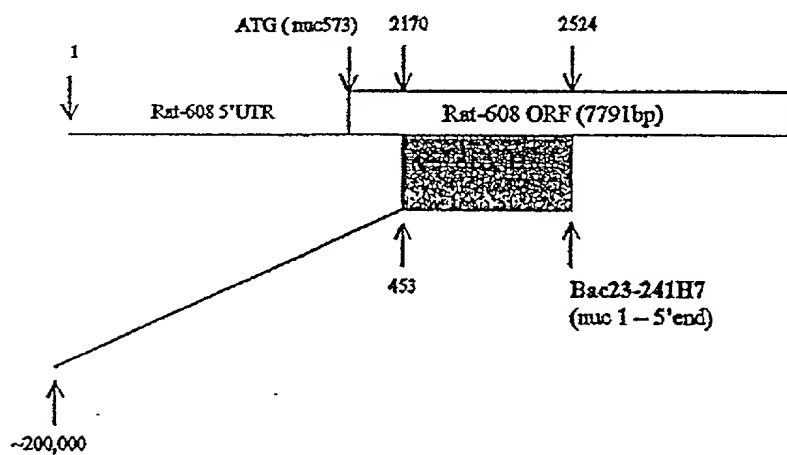


Figure 7

Nuc 1

↓
TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTGCATTTA
TCCAATGGAATACTACTCAACTTATTAACCAACGACTTTCATAAAATTTTAGGCAAAT
GNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACACTCATGGTATGC
ACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAGGGAAGAACACATAGC
TTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGTTTGGCTTATGAGTGAATCT
TGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAGTGTAGGAAGAAGTAGTTAATCA

↓ Nuc 390

AGGTGACAAAGTATCCTGACACATTGGGAAAAGACCACAGTCCAGGAAAAGTGTCTTAA
GGATTTCATATTAACCTCCAGTTTCCCATGTGCAGCTCTGAGACTTTGGCAGATCAGACACTT
AACTTACCAGCTTCTACACAGAGCAGTTACTATCCTTGCATTCACAGATGGAGTGTGA
CCATTAAAGCTGCTGACCAAAACATGAGTCTGACTTGITAATAATCTTAAATAACAAATGTGT
TGTAAGATATGTGACCAAAAGAGCATGGTCACTGCTATTAACTTTGATGTTCTATGGACTCT
TAATTTTATGGTAGAAATGTCAACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAAGAG
GATGGCTTTCAGTTTGTAAAGTAAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCA
ATTTAAACCCCAAGTAGGCATAATTTTAATGCTTACTTCATCAGAAATATGCTAATTTGAC
TCTTCTAAAAAGACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAAT
AAATAGGAGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTT
ATAGGTTATATAAGAGGTGGGAAAAGGAATATTGAAACATCTAGAAGTAAATGTATCCTG
AGTAGCGATCCTGGGAAAATACGTAAGTCTAACACACTGCAATCATCTCTGTGTTTGGT
GGAGCTGAGGTCTGGAAGGCTGACCTTGGTTAGAAATAACCTACCGAATACAGAGCTAT
GATGCCATATTTCAGGCAAGTGTCCACATGCGGACATCTTAAAGAAATATGGTTGTCTCTG
CAGTGTCTAAGGACCTTGTCTGTCACACAGGTCTCCAGGTTAGTGTCTAAGTCTGATGCT
TTGACTCTTAAATTTCTACCTTGATCATTAAATGACTAGAAATCACTTGGTGTATTAGCAACTG
GATATGGAATATTAATAATTTGTACCCAAAGCCAGGCCACCTCAGCTTTGGCAGCTCCATT
ATTGCTGGAGGCCAGTCACTGGGTTTGAATCAACTGTACTGTTTCTACTTACAAGACGC
ATTACCTGAGATGAGTCAATTTTCTTCAACAAGTCTTTTGAAGAGTCAATTAAGACATATT
CTGATGAAGTAAAGCATATAAAGTGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGA
TGCTATTATAATGTGGAATAAACTGACTTTAAAAAAGGCTTATGATACTTGTGACA
GAGTAAATCTTCCATAAATATCATCTGCAATTTATAAATTTTTCATAATCCATCAATTA
AAACCTTTAGAAATTTTGTAAACACAAAGATCCCTAGGCCCTGCCCCTAGGATGGTCTGTA
TGGTGGGCTGAGAGATGGAGCTTAAAGAACTTACTGTCTCAGGAGCAGATCTTCAGAAC
ATCTGGCTCAAAACATTTATCCAAATGCTCATCAAAAGGCTCACTCAGATGTGCTTCAACC
ACAGGGATTAAACAGTCATTTTGTGACATTTCTCAAAAGGCTGGAAGGCTGTAGAGGAA
CAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTGATAACCTCAGCACAACA
AACAACAACGACAACAAACCCATAAAATTTATCAAGGCAAAATGAACTAAGCCATATGCA
GAAAAAGTATTATATGTTCTCATTTGTGGGGTGTTTTCTTAAATAGTCAAAATATGCAGAA
ATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTTGTAGATTAGAGGACA
CAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAATGTATAATATGAGGACTGTA
TTAATAATTCTATTAAAGATACACAGCAAACGAGTGTATCTTACTAACACACACACTTAC
ATAGAGAGAATAAAGTGATAGATACGTTTGTATTATCTTCATGTAGCTGATAATTTTCATAT
TGTACACCTCAACATAGATAACCAACAAGAGGAAGAGGATAGGTGCTCTCCAGGGC
GGAAGAGTACATTGGAAGTCAAGACACCAATTGTGTAGATGTACCACATGGAGGAGCTAGA
GAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCCTATAGGTGGAACAACATTTATGAGCT
AACCAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATCAAAAGATGGCCTAATCGG
CCATCACTGGAAGAGAGGCCCATTTGACTTGCAAACTTTATATGCCCCAGTACAGGGGA
ATACCAGGGCCAAAAAGGGGAGTGGGTGGGACGGGAGTGGGGGTGGGTGGATATGGG
GGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATACTAATAAAAAATGAAAAA

↑ Nuc 2317

AAAAAAAAAAAAAAAAAAAAAGGAAGGTGAGACACCTCACTTCACTGCTATCTCAACTT
GCAACAGAAAGGGAGTCACAAACCCAGGACAAACACAGTGATTGAAGCGTCTTTGAA
TGTTATTGCTGTTGTTGTTACCAATCATTAGCATATATTCATTGTGAAAACCTTACGGGGT
CTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGCACCCTACCA

0991630.030402

TGTTAAATACAGTTTGTAGGCTTAAGTGTACGGGAACCTCATGTGGTATTTCATTTAACGGCTCT
 CTCTCTTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTCTCCAGCTACCTTGTAC
 CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATCAAGCAGAGAGTATT
 CAAAATATTTTCAGTGATGTAAACAGTGACAGCTGAGGCATAGAAGTAATCATGATGAATCT
 TTAATTTGGGTTTAAACTCTATTCTATAACAGCTCCAGGTTGGGAGGGGATCTAGTGACCTTTCG
 CACGTGGGGTTAAAGATATTTTCTAAACAAGAGAAGCAGAATTCTCTCTTGGCCATGCTCC
 CCATCACTGTGTCAAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAGAGCAGACAGTGT
 TGCCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGCTCAGTTTACTGTCTCCCGGTCAAT
 AGTTGGCTCTGAAAAGGCCCATGTGTCCATATTGGCAAGGACTTGCAGACATGCTAGAAA
 GAAATTTGACCTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTATTTTGGAAAGGTTAAG
 CCAATATATAAAACACATATTAATAATACAATGTTACAAAATTTGATCATATAAAGAA
 GTACATTCTATAAATGCAATTGTGAAAAATATATATAATTTTTATCTATTTACTGGTGCAAAG
 TTTTCTAAATTTGCACATGTACTATTTTTATATTATAAAAAATATTTTTAAAAATGTATATAAA
 AGTGTA AAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAAATTTAATTTGGTCCGAT
 AACATTATTATGATCTCTAATGACAGGGATCTGCTTTTCATTGGGAAATGAGAAGCTATG
 AAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCAATGGGAAGCTAGCACAC
 ACTGGTTTATAAAGAGAACAGTTTCTCTGAGTCTATGCAAGTTTACACTCTAGGAATAAG
 AGTTCTCTTTCTCCAGATTCTACTAGCATTCTTTGTTGTCTCATTTATCTCTTGGATGATGAG
 CATTATAAGTGGAAATAAGATAGGATCTCAAAGGAATGTCAATTTGGATGGCCGTGAACAAT
 CTTTACAGGCTCTTTCTTTCAAGTTCACTAGTCTATTTCATTTATTGGATAATTGGGGGATGGTGT
 TAATTTTTTTGCAGTTCTTATGGAATTCCAAAAAACAAAAACAAACAAACAAACAAAA
 ACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACA
 GAATTTATTGCTACATGTTTCATTATTTCACGACAGCCAAGAATGTGGAACCAACTTACGTA
 GCGCTCAAAATATGAAACGGATATAAGAAAATGTGGAATGTGTACAACAGAGATCCCATGTG
 GGCATAAAAGAGTGAATTCATGACATATGACAGAAATGGATGCAACTGGAATCAATTG
 GGCTAATCAAAACAAGACAGACTCAAAAGGAAACACCGGTGAGCTTCTCTGACAAACA
 GAAGCTAGATTACACITGTACGTGGCATGTGTGTTTAGAATTTTATTTFAGTTATACACT
 ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAAACAAGCTAGCTGGGG
 TGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAACATAGGTG
 CGAGGATGA AATGCAATTAATTGTATGCTAACTAAACCACAGACAGGAGGCCACAGTTTCA
 AACCAGGGTGAAATCCACGACAGAGAAAGGGGAAGTAGACACAAAAGTTTCCCACTAAC
 CAAGAAGCCATTTCAGTTTGTCTGCTGCTGGGAGGGGGCTTCCAGTTTCTCCAGTCTGAC
 ACTGTGTATAACAACCAAGTTTGACAATACAAAGTTGGCATGATGGAATGGTTTTTGTGCTATT
 TTTCAATTTTTTTCTTACTGTTTTTGTGTTGTGTTGTTGTTGTTGGTGGTGGCTGTGTTTTC
 ATTGTGTTCTTTGAGAGAGAGAAGAAACATGA AATTTGGTGGTGGTAGGAAGCTGGAACCG
 ATCTGGAAGAAAGTTGGGGAAAGAGAAAATTTGATGAGCATATTTAAACAAACAAACA
 AACAAACAAAAGGTTCAATTTGCCACAAAAGGTTGTGAATTA AATTAACCAAGTTACGACT
 CTAAAGAAAATATTCCCAATTATTCACAGAGTTGCTATGTATGCTGTGCTAGGACTTTG

0954530.030462

TGTGCGCTGCAACATCTTGAAAACATTGCGGTACATGATGGAAGCCAGGCACAAAAAGCC
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAAATACATGGATTCTTAGCAAACAGA
GTAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAATACTATTGATGTGGAATG
ATCTTAATTTGGGGAAAAGACAATTTCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
ATCCCTGTGGATATTGTAATAAACAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
GCAGGAGGAAAGTGAAATGAGATGGTAGAAAAGGAAAGTCATATAACCATGGCTTCTCTGCTG
GGTGGAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGGAAGGATC
AAAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAAGTTCATGGTCAA
TGTGTGTGGAGACACCATAATAAACTCCTTTTTTGTGCTAACTAAAAACCACTAAAAATC
TAAAAACAAAACATTTTTGCACAAGAATTATTATTATTCAATAAAGATGTTTAAATGGGG
GAAGTTGAAGTTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTACGTTT
TTTATTATTAATTACTCTGTCTGTCATTATTATCATCATCATTATCGTCATCATCACTA
ATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACITAGTGAGGCAATCTT
GGAAGAGGGAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGAA
ATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCATAGTCCACTACATTACTTTG
TATTCTAATATTAAGAAAATAATAAACCCATTTCTGTGCACTTATCACCCAGGCTCAACAG
TTATCTTGGCCACAGATCCTGTCTCACTGCATCCTGTGCCACCTGAGTCCACTTAGCGTTCTG
AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA
AAGTAACCATGACATCTCTCTGTTTCTTTAGCGATAAGTTCTTAATATTATATGGCTGCTTGT
GTATGTTCTAATTTCTCTAATATTGTTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGT
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTCTATATTTATAAACTTTCTCTCT

Exon 4 ? (Nuc 11286) ↓

TTGTGAAAGTTCCTTTGAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGAATCAGGAGCTTGACCCGG
TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTACGGACTCACCTTGC
TCCGCTTGCTACATCTAGAAAGAAACCGGCTGACAAAAGCTCCATCCAGACACATTTGTCTCTT
GAGCTATCTCCAGATATTTAAAACTCCTTCATTAAGNAOCTGTACTTGTATGATAACTTCATTG
ACCTCCCTCCAAAAGAAATGGTCTCTCTATGCCAAACCTAGAAAGCCTTTACTTGCAATGGAA
ACCCATGGACCTGTGACTGCCATTTAAAGTGGTTGTCCGAGTGGATGCAGGGAAAAACAGGTA

Nuc 11680 ↑

ACTATCTGTGTGTGTGTCTTTTTTATARKACGTTATTTCTCAATTTTCAATTTAGAAATGA
TATCCAAAAGTCCCCATAAOCCTCCCCCCTTCCCTACCTACCCATTCCCATTTTTTGG
CCCTGGCATTCCCTGTACTGGGGCATATAAAGTTTGGGTGTCCAATGGACCTCTCTTCC
AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
TACTGGTTAGTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967 ↑

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

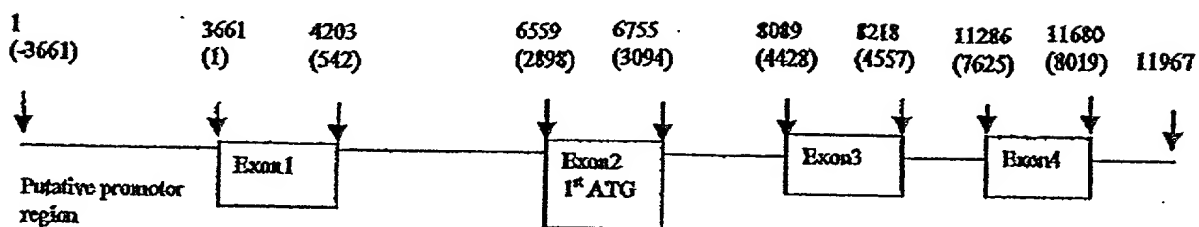


Figure 10

cDNA_rat	1	ogagagacgacagaagggttaoggctgcgagaagacgacagaaggggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgtctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgaactgaggccaggcaaaaacgogcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggacccctttcatagacactgatgacacgtttaogcaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaacgctgggccttcggaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaaatggatgatagcattaaaataacagaagcgctccagg
genomic_hu	1	-----
cDNA_rat	401	tototgaagottoagtcccccagctgaagocagaaaagaactaagoccaa
genomic_hu	1	-----
cDNA_rat	451	taagccttttgatccctttggaagcaaagaactttccttccttggggtga
genomic_hu	1	-----
cDNA_rat	501	agactotcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcagggtgagaggcagggaaagtcag
genomic_hu	1	-----
cDNA_rat	601	oggcttggtgatotccctoaotgotgtctgootgggtgtoacccctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcctgtcctgcgcgtgtgcctgctatgtgccacagaggtgcac
genomic_hu	1	-----
cDNA_rat	701	tgtacatttcggtaacctgacctccatcccagatggcatcccgccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaaatttaggatataacagccttactagattgacagaaaacg
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaaactggagttactcatgtgtgcacagtaatggc
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcagtccttgaggt
genomic_hu	1	-----
cDNA_rat	901	cttaaaaatgagctataaacaagtcacaatoattoggaaggatactttct
genomic_hu	1	-----

cDNA_rat	951	acggactcgggagcttggtccggttgacactggatcacaacaacattgaa
genomic_hu	1	-----
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu	1	-----
cDNA_rat	1051	tttagaaggaaaccggctcacaagctocatccagacacatttgtotcat
genomic_hu	1	-----
cDNA_rat	1101	taagctatctccagatatttaaacctctttcattaagtacctgttcttg
genomic_hu	1	-----
cDNA_rat	1151	tctgataacttcctgacctccctcccaaaagaaatgggtctcctacatgcc
genomic_hu	1	-----
cDNA_rat	1201	aaacctagaaagcctgtatttgcattggaacccatggacctgtgactgcc
genomic_hu	1	-----
cDNA_rat	1251	atttaaagtgtgtgtgtgagtgatgcagggaacccagatataataaaa
genomic_hu	1	-----
cDNA_rat	1301	tgcaagaagacagaagctcttcagtcctcagcaatgtcccctttgcat
genomic_hu	1	-----
cDNA_rat	1351	gaaccccaggatctctaaaggcagacccctttgtatggtacocatggag
genomic_hu	1	-----
cDNA_rat	1401	ctttcctatgtacaaagccaaccattgatccatcactgaagtoaaagagc
genomic_hu	1	-----
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	-----
cDNA_rat	1501	catagaacocctttggctccttgtctttgaacatgacanannntctggaa
genomic_hu	1	-----
cDNA_rat	1551	ataaggccgacatggtctgtagtatocaaaagccatcaaggacatcacca
genomic_hu	1	-----
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgctaaatgcgtcatttct
genomic_hu	1	-----
cDNA_rat	1651	cacaaatcttgtgtgcagtgtagattataatcacatccagccagtgtggc
genomic_hu	1	-----
cDNA_rat	1701	aaottctggctttatacagtgactctcctotgatactagaaggaagccc
genomic_hu	1	-----
cDNA_rat	1751	cagcttacggagactccttcactgtcttctagatataaacagggtggctot
genomic_hu	1	-----
cDNA_rat	1801	taggcctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	-----
cDNA_rat	1851	ctttttggttccaacaagaaaaattgtcttgcagctgaacagaactgcc
genomic_hu	1	-----
cDNA_rat	1901	accacacttagoacattacagatocagttttccactgatgtcctaatcgc
genomic_hu	1	-----
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaatggaocatgatoc
genomic_hu	1	-----

cDNA_rat	2001	tgatgatgaacaatoccaaactggaacgcactgtctctgggtggcggaact
genomic_hu	1	-----
cDNA_rat	2051	attgccotgagctgtccaggcaaaaggogacccttcacctcacttggaaatg
genomic_hu	1	-----
cDNA_rat	2101	gottotagctgatgggagtaaagtgagagcccttacgttagogaggatg
genomic_hu	1	-----
cDNA_rat	2151	ggcgaatcctaataagacaaaaatggyaagtggaaactgcagatggctgac
genomic_hu	1	-----
cDNA_rat	2201	agctttgatgcaggtctttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	-----
cDNA_rat	2251	ggatgtttcacatacaggataactgtggtagagccctatggagaaagca
genomic_hu	1	-----
cDNA_rat	2301	cacatgacagtggagtcacgacacagtggttacgggtgagacgctcgac
genomic_hu	1	-----
cDNA_rat	2351	cttccatgcctttccacgggtgtccagatgottotattagctggattct
genomic_hu	1	-----
cDNA_rat	2401	tccagggaacactgtgtttctctcagocataagagacaggcaaattctta
genomic_hu	1	-----
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgccaaaagatoaaggtcac
genomic_hu	1	-----
cDNA_rat	2501	taccaatgtgtggctgccaaaccatoaggggcogacttttccagttttaa
genomic_hu	1	-----
cDNA_rat	2551	agtttcagttcaaaaagaaaggccaaaggatggttgagcatgacagggagg
genomic_hu	1	-----
cDNA_rat	2601	caggtggatctggacttgagaaacccaactccagtgtttcccttaagcag
genomic_hu	1	-----
cDNA_rat	2651	ccagcatctttgaaactctctgcatcagctttgacagggtcagaggctgg
genomic_hu	1	-----
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaacatagagacttaatac
genomic_hu	1	-----
cDNA_rat	2751	atcgccggcggtggggattccacgctccggcgattcaggagcataggagg
genomic_hu	1	-----
cDNA_rat	2801	cagctccctctctctgctcggagaattgaccgcgaacgctgggcagcact
genomic_hu	1	-----
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaagcaagaaaatacca
genomic_hu	1	-----
cDNA_rat	2901	cagtaaagccagtgcactgggtgttccctcgtggaactcactgaogag
genomic_hu	1	-----
cDNA_rat	2951	gaaaaggatgcctctggoatgattctccagatgaagaattcatggttct
genomic_hu	1	-----

cDNA_rat	3001	gaaaactaagggttctggtgtccaggaaggtcaccaactgctgactotg
genomic_hu	1	-----
cDNA_rat	3051	gaccagtaaatcatggttttatgaagagtatagcttctggcxcagaagtc
genomic_hu	1	-----
cDNA_rat	3101	tcaactgtgaatocacaaacactacaatctgagcaccttootgatttcaa
genomic_hu	1	-----
cDNA_rat	3151	attattttagtgaacaaacggtacagctgtgacaaagagtatgaaccocat
genomic_hu	1	-----
cDNA_rat	3201	ccatagcaagcaaaatagaagatacaacccaacccaaacccaatcattatc
genomic_hu	1	-----
cDNA_rat	3251	tttccatcagtagctgaaattogagattotgctcaggaagggaagagcatc
genomic_hu	1	-----
cDNA_rat	3301	ttcccaagtgacacacootgtaacagggggaacatggtacacctatggcc
genomic_hu	1	-----
cDNA_rat	3351	ataccaacacatatagtagctttaccagcaaaagccagtagctcttgag
genomic_hu	1	-----
cDNA_rat	3401	ccaataaatccaacagaaagttatggacctcagataacctattacaggagt
genomic_hu	1	-----
cDNA_rat	3451	cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cDNA_rat	3501	gcttctccagtoacccttcagggttcacacaccactgcctogtctttattt
genomic_hu	1	-----
cDNA_rat	3551	cacattccctagaaacaacaatacaggttaacttccootgtgccaggaactt
genomic_hu	1	-----
cDNA_rat	3601	gggaagagagaggacaatttggagcagagggagaggttaaaaaccacata
genomic_hu	1	-----
cDNA_rat	3651	gaaccccagttctccagcggcatagacacaggaactgtgaggccagcaatc
genomic_hu	1	-----
cDNA_rat	3701	aagggaactgctaacaaaaatgtgagccaagttccagccacagagtaccc
genomic_hu	1	-----
cDNA_rat	3751	tgggatgtgccacacatgtccttccgcagagggggtcacagtggtactg
genomic_hu	1	-----
cDNA_rat	3801	cagcaactgtcagttccaagttcatcccaagtgccctcccaaaaactaat
genomic_hu	1	-----
cDNA_rat	3851	aatgttggggtcatagcagaagagtctaccactgtggtcaagaaacoact
genomic_hu	1	-----
cDNA_rat	3901	gttactatttaaggacaacaaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cDNA_rat	3951	ctacaaaatattccggaggggaaagtaaccaagtgattoctacggaagca
genomic_hu	1	-----
cDNA_rat	4001	agcatgaactctgctccaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cDNA_rat	4051	caatagtggtcacctgagcatgootgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cDNA_rat	4101	cagtggaaacaacaccacttcccagccccctcagcacacctcaatacca
genomic_hu	1	-----

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

Figure 12

ATGAAGGTAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCCTTGGAA
GGAAATCAGCTCACTAAGCTCCACCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACCTTGTCTGATAACTTCTTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCCTCAGCAGTGT
CCACTTTGCATGAACCCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG
TGTGCCAAGCCAACCATTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT
GCTTTTCATCTCTCCCAAGGTTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTCACTGAA
GAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTACGGTCAC
ATTGAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCTCTGATACTAGAAAGGAGCCACTTG
CTTAGTGAAACACCCGAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACC
AACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAAC
AGAAGTCCACCACATTTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA
GCAGAGATGAGGCCAGTGAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT
ACTGCTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCACCCACACGTGGAT
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCCTTATGTCAGTGAGGATGGACGGATCCTAATAGAC
AAAAGTGGAAAATTGGAATCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATAAGCAGC
AATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTGAAGCCTATCAG
GAAAATGGGATTTCATCACACAGTTTTCATTGGTGAACACTTGATCTTCCATGCCATTCTACTGGTATC
CCAGATGCCTCTATTAGCTGGGTATTTCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA
GTTCTAAACAATGGCACATTAAGAATATTACAGGTCAACCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAAACCATCAGGGGTTGATTTTTTGTATTTTCAAGTTTCAGTCAAGATGAAAGGACAAAGGCC
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTTAAGGAG
CCACCAGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA
AGTAAGAGGCACAACATATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG
GAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACGTGTGGAG
AAAGCTAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCACCCCACTGGTCAAC
CAACTCCCAAACATACCTGGTGGAAGAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTTATG
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGAT
AGTCTATGACAAACATAAATTATGGCACAGAATTCTCTCTGTTGTGAATTCACAAATACTACCACCT
GAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAACCTACAGCCATGTCAAAGAATATAAACCCA
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCACTCACTGTCTTTCCACTGCTACTTGGGA
GCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCAATA
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAATGCTTAGTAGCACCACCAACAACTATTATTA
GAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC
TTCTATTCTCAGACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACAGCTGCTCAT
TCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAAACATCCCGCTGTTTCAGACGCTTTGGGAGGCAGAGG
AAAATTGGCGGAAGGGGGCGGATTATCAGCCATATAGAATCCAGTTCTGCGACGGCATAGATACAGC
ATTTTCAGGTCAACAACAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT
GTGACATGCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCTCTAGTC
CAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAAACAACCCCAATAAAATATTTTC
AGGACTGAAATTTCCCAAGTGACTCCAACCTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG
GAAAAAATCACAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACAATGAAGCTAAAAGAGATTCA
GTGATTACATCGTCACTTTTCAGGTGCTATCACAAGCCACCAATGACTATTATAGCCATTACAAGGTTT
TCAAGAAGGAAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAAGGAAT
CAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTTACCACAG
AGACAAAGTTTCCCTTTCCATTTTACCACACTTTTCAACAAGTGTGATGCAAAATTCATCTAATACCTTG
ACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCAACAAAGAAGGAGCTTCCC
TTCCCACCCCTTAACCTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATCATATCA
ACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCACTTATGAAACCCAAACA
GAGAGATCTAGAGCACAACAATACAAAGAGAACAGGAGCCTCAAAGAAGAACAGGACTGACCCAAAC
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACT

CATTCCCCACCAGAAAAACACAACCTGGGATTTCAAGCACAAATCAGTTTTTCATTCAAGAACTCTTAATCTG
ACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCTGAAACA
ACTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCATACCACCA
TTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACT
GACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGG
CACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGGGCAGAA
TACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGCATGTTG
GCTACTACAGGCCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAAGAACAAGAAG
AGTGACTTTGATAAGAAACCAGTTCAGAAGCAACAACCTCCAACTCCTTCCCTTTGACTCTTTGTCT
AGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTACTATTCCAGCTAATCTCA
GATGCCCTTTCTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCATTCATTGGACCAGAGTTTCAGGA
CTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCAATGGTACCCTGTCCATCCAG
AGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTT
CATGTACACCTTGTCTGTGGTTTCTATCCTCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTT
CATTCCCGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCTGGATT
CTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGA
ACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCAGGT
GGCCAGGATTCAGTCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGG
CAAGTCATTGTAGGCACCTTGGGGTGAAGTTTAAACTGCCCTGTACTGCAAAAGGAACCTCCTCAGCCC
AGCGTTTACTGGGTCTCTCTGATGGCACTGAAGTGAACCATTACAGTTTACCAATTCCAAGTTGTTT
TTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATT
GCTACCAGTTCCACTGGTTCCGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCC
AGGATAGAAGCTGCATCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAAGTGTCTCA
GCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCCAGCAGTGGAGC
TGGATCCACGTCTACCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTAC
TTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCT
GCCAAAATTGACCACAAGCAGTATTTTAGAAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGC
AAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCA
ATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTTTCAACAATGGAACCTTTATCTTC
AACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACCTTGCTATGCCAGAACACCCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGATAAGGCAGAGTAACAAAACCAACAAGAGA
ATCAAAGCTGGAGACACAGCTGTCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAATATTTTGG
TTGTGCTCTTCCAATGACATGATTTCTTCTCCATTGATAGGTACACATTTTATGCCAATGGGTCTTTG
ACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCAGTGGGGAT
GACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGA
ACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAACACTTTGACTGCAGAGCTGAAGGGACACCA
TCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCTCAGACCCCATACTATGGAAGCAGAATC
ACAGTCCATAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTTCAAGTTTCAAGCCGACTTTATCTGT
GTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGA
CCGACATTTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATTGC
TCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCCAATGGA
CCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTCTTAAACAACTCGGGAGGAT
GCAGGAAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAAT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCA
CTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTA
GACAGGCCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACA
GCTTATGACAGAGGAACTATATCTGTAAGGCTCAAAATAGTGTGGTGCATACACTGATTACTGTTCCA
GTAATGATTGTAGCCTACCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACAGGG
GCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATGCCTGAC
CACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAAGGTACC
CTAGTCATTTCAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCACCTTGGT
AGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cDNA	CGAGAGACGACAGAGGTTACGGCTGCGAGAAACGACAGAGGGTCCAGAAAAAGAAA
human 5+3 corrected	-----
mus_cDNA_5	-----CAGAGGGTCCAGGAAA-GGAAA
rat_cDNA	GTGCTGGAGGGGAGTGGGACAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG
human 5+3 corrected	-----
mus_cDNA_5	GTACTGGAGGGGAGTGGGACAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAA
rat_cDNA	GCCAGGCAAAAGCGCGGGAAGGATTTTGTGTAGCTTGGGACCTTTCATAGACTGAT
human 5+3 corrected	-----
mus_cDNA_5	GCCAGGCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAAGACTGAT
rat_cDNA	GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCTGGGCTTCGGAAAG---GA
human 5+3 corrected	-----
mus_cDNA_5	GACGTGTTTGTGCAAAAGCAAGCAATTTGAGGAGAAACGCTGGGACGTTCGGAAAGAAGGA
rat_cDNA	GTGATTGATTAGTACTTGCAGTTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA
human 5+3 corrected	-----
mus_cDNA_5	GTGATCGATTAGTACTTGTAACTTAGGTGAGTTT---GAGAACTAACTAAGCTATATACTA
rat_cDNA	TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCAGCAGCAGCAGGAAGCTTTGTTAAT
human 5+3 corrected	-----
mus_cDNA_5	TTGAGGGAGGAAGAGAGCATT-----CCAGCAGCAGCAGGAAGCTTTGTTAAT
rat_cDNA	TTGGAATGGATGATAGCATTAAATAACAGAGCGGCTCCAGGTCTCTGAAGCTTCAGT
human 5+3 corrected	-----
mus_cDNA_5	TTGGAATGTATGATACCATTAATAACAGAGCGGCTCCAGTTCTCTGAAGAGTCAGT
rat_cDNA	CCCCCAGCTGAAAGCCAGAAAGACTAAGCCCACTAAGCCTTTTGTATCCCTTTGGAAGCA
human 5+3 corrected	-----
mus_cDNA_5	CCCCCAGCT-----A-GTGTAAGCCTACTAAGCCTTTTGTATCCCTTTGGAAGCA
rat_cDNA	AAGAACTTTCCTTCCCTGGGGTGAAGACTCTCTCAGAGATTTCCTGTCTCTGCCTATG
human 5+3 corrected	-----
mus_cDNA_5	AAGAACGTTTCCTTCAATCAGGTGAAGGCTCTCTCAGAGATTTCCTGTCTCTGCCTATG
rat_cDNA	TTACAAGAGGAATCAAAACAGACAGAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAG
human 5+3 corrected	-----ATGAAGGTAAAGGCAGAGGAA
mus_cDNA_5	TTACAAGAGGATTCAAAAGCAGACAGAGAGCTCAGGATGCAGAGAGAGGCAGGGAAG
	*** ** * ***** *
rat_cDNA	TCAGCGGCTTGTGATCTCCCTCACTGCTGTGTGCTGGTGGTCACTCCCTGGGAGCAGGG
human 5+3 corrected	-----
mus_cDNA_5	TCAGCTGCTTGTGATCTCCCTCACTGCTGTGTGCTGGTGGTCACTCCCTGGGAGCAGGG
	*** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA	CCTGTCTCGCGGCTGTGCTGTGTATGTGCCCACAGAGGTGCCTGTACATTTGGGTACC
human 5+3 corrected	-----
mus_cDNA_5	TCGTCTCTCGCGGCTGTGCTGTGTATGTGCCCACAGAGGTGCCTGTACATTTGGGTACC
	***** ***** * ***** * ***** * ***** * ***** *
rat_cDNA	TGACCTCCATCCAGATGG-CATCCCGGCCAATGTGGAAAGCAATAAATTAGGATATAAC
human 5+3 corrected	-----
mus_cDNA_5	TGACCTCCATCCAGAGCGGCAATCCAGCCATGTGGAAAGCAATAAATTAGGATATAAC
	*** ***** * ***** ***** * ***** * ***** * ***** *

00041630-030416

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGCCTTACTAGATTGACAGAAAACGACITTTGATGGCCTGAGCAAACTGGAGTTACTCATG
AGCTTGGTTAGATTGATGGAAACAGATTTTCTGGCCTGACCAAACTGGAGTTACTCATG
AGCCTCCTAGATTGACAGAAAATGACITTTTCTGGCCTGAGCAAACTGGAGTTACTCATG
*** * ***** ** ** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGCACAGTAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
CTTCACAGCAATGGCATTACACAAATCCTGACAAAGACCTTCTCAGATTGCAAGCCTTG
CTGCACAGCAATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGGAAGGATACTTTCTACCGA
CAGGTCTTAAAAATGAGCTATAACAAAGTCCGAAAGTCCGAAAGGATACTTTTATGGC
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGAGAAGGATACTTTGTATGGA
***** ** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCGGGAGCTTGGTCCGGTTGCACTGGATCACAACAACATTGAATTCATCACCCTGAG
CTCAGGAGCTTGACAGGATTGCACTGGACACACAAATATGAGTTTATAAACCCAGAG
CTCAGGAGCTTGACCCGGTTGCACTGGATCACAACAACATTGAGTTTATCAACCCGAG
*** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCGTTTATGGACTTACCTCGCTCCGCTTGGTACATTAGAAGGAAACCGGCTCACAAG
GTTTATGGGCTCACTTTCTCCGCTGGTGCCTTGGAGGAAATCAGCTCACTAAG
GCGTTTATGGACTTACCTCGCTCCGCTTGGTACATTAGAAGGAAACCGGCTCACAAG
* ***** ** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCCATCCAGACACATTGTCTCAATTAAGCTATCTCCAGATATTTAAACCTCTTTTCAAT
CTCCATCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCAAT
CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAACCTCTTTTCAAT
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTACCTGTTCTTGTCTGATAACTTCCT-GACCTCCCTCCCAAAGAAATGGTCTCCTA
AAGTCTCTATCTTGTCTGATAACTTCCT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
AAGTACCTGTTCTTGTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTC
*** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CATGCAAAACCTAGAAAGCCTGTATTTGCATGGAACCCATGGACCTGTGACTGCCATTT
TATGCTTGACCTAGACAGCCTTTACCTGCATGGAACCCATGGACCTGTGATTGCCATTT
TATGCCAAACCTAGAAAGCCTTTACTTGCATGGAACCCATGGACCTGTGACTGCCATTT
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAGTGGTTGTCTGAGTGGATCCAGGGAAACCCAGATATAATAAATGCRAGAAAGACAG
AAAGTGGTTGTCTGAGTGGATCCAGGGAAACCCAGATATAATAAATGCAAAAGATAG
AAAGTGGTTGTCCGAGTGGATCCAGGGAAACCCAGGTA-ACTATCTTGT---TTGTTTG
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGCTCTTCCAGTCCCTCAGCATGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
AAGTCCCTCTAGTGTCTCAGCAGTGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
TTTCTTTTATARKAGTATTTCTCAATTTCAATTTAGATGATATCCCAAAGTC-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACCCCTTGTCTATGTTACCATCTGGAGCTTCTCTATGTAACAAGCCAAACCTTGATCCAT
AGCCGTTAGCTATGTTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAAACCTTGATCCAT
-CCCATAACTCCCAAGCA-CTTCCCTACCTACCATTC-CAATTTTGTGGC
** * * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGTGAAGTCAAGAGCCTGGTTACTCAGGAGGCAATGGATCTGCTCCACCTCACTC
CCTGAATCAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTCTATCTCTCCCC
CCTGGCATTCCTC-
* * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGATTTCTAGAAACCTTTGGCTCCTTGTCTTTGAACATGACANNTNTCTGGAATA
AAGTTTCTAGGCAACCTTTGGCTCCTCCTTGAATATGACAGATCAGTCTGGAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGGCCGACATGGTCTGTAGTATCCAAAGCCATCAAGGACATCACCAACTGCATTCATG
AAGCTAACATGGTCTGAGTATTCAAAGCCCTCAAGGACATCACCAACTGCATTCATG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

AAGAAATGACTACATCATGCTAAATGGTCATTTTCCACAAATCTTGTGTCCAGTGTAG
AAGAAATGACTACATCGTGCTAAATACCTTCATTTTCAACATTTTGGTGTGCAACATAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

ATTATAATCACATCCAGCCAGTGTGGCACTTCTGGCTTTATACAGTGACTCTCTCTGA
ATTACGGTCACATTCAGCCAGTGTGGCAATTTTGGCTTTGTACAGTGATTCTCTCTGA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG
TACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGCTCTTAGGCTGAAGACATTTTACCAACATAGAGGCTGATGTCTAGAGCAGACCTT
TGGCTCCTAAGCTGAAGACATTTTACCAACATAGAGGCTGATGTCTAGAGCAGATCCCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TTTGGTTCCACAAAGAAAAATTTGTCTGCAGCTGAACAGAACTGCCACACACTTAGCA
CTTGGTTAATGCAAGACCAAAATTTCTTGCAGCTGAACAGAACTGCCACACACTTAGTA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CATTACAGATCCAGTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGAGCAGAGATGAGGC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGAGAGACTCAAATGGACCATGATCTGATGATGAACAATCCAACTGGAAACGCACTG
CAGTGAAACACAAATGGACTATGATTTCAAGGGATACAAATAGTAAGCTGGAAACATACTG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TCTTGGTTGGCGGCACTATTGCGCTGAGCTGTCCAGGCAAGGGGACCTTCACTCACT
TCTTGGTAGGTGGAAACCTTGGCTGAACTGCCAGGCGAAGGAGACCCACCCACAGG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCCTTACGTTAGGAGGATGGGC
TGGATTGGCTTCTAGCTGATGGAGTAAAGTGAGAGCCCCCTTATGTCACTGAGGATGGAC

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGATGAG
GGATCCTAATAGACAAAAGTGGAAAAATGGAACTCCAGATGGCTGATAGTTTTCACACAG

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GTCTTTACCACTGCATAAGCAACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA
GCGTATATCACTGTATAAGCAGCAATATGATGATGCAGATATTCTCACCTATAGGATAA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
CTGTGGTAGAACCTTTGGTCCGAAGCCTATCAGGAAATGGGATTCATCACACAGTTTCA

rat_cDNA
human 5+3 corrected
mus_cDNA_5

CGGGTGAGAGCGCTGACCTTCCATGCTTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAAACACTTGATCTTCCATGCAATCTACTGGTATCCAGATGCCCTATTAGCT

rat_cDNA
human 5+3 corrected
mus_cDNA_5

GGATTCTTCCAGGGAACACTGTGTCTCTCAGCCATCAGAGACAGGCAATTTCTAACA
GGGTATTCCAGGAAACATGTGCTCTATCAGTCATCAGAGACAGAAAGTTCTAACA

09501630-030402

rat_cDNA human 5+3 corrected mus_cDNA_5	ATGGGACCTTAAAGATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG ATGGCACATTAAAGATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTGG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCCAACCCATCAGGGGCGGACTTTTCCAGTTTTAAAGTTTCAGTTCAAAGAAAGGGCC CAGCCAACCCATCAGGGGTTGATTTTTGATTTTCCAGTTTCAGTCAAGATGAAGGAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAAGGATGCTTGAGCATGACAGGGAGGCAGGTGGATCTGGACTTGAGAAACCAACTCCA AAAGGCCCTTGAGCATGATGGAGAACAGAGGGATCTGGACTTGATGAGTCCAAATCCTA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTCCTTAAGCAGCCAGCATCTTTGAAACTCTCTGCATCAGCTTTGACAGGGTCAG TTGCTCATCTTAAGGAGCCACAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC AGGTTGGAAAACAACCTCAAGCACAGTAAGAGGCACAACTATCGGGAAATTAACACTCC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GGCGGCGTGGGGCAITCCACGCTCCGGGCAITCAGGGAGCATAGGAGGCAGCTCCCTCTCT AGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGGAGGCATTTCCTCCCT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGCTCGGAGAAATGACCCGCAACGCTGGGCGACCTTCTAGAAAAGCCAAAAGAAAT CTGCTAGGAGAAATGACCCACAACTTGGGCGGCACTGTTGGAGAAAGCTAAAAGAAATG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGCCAAAAAGCAAGAAATACCACTAAGCCAGTGGCACTGGCTGTTCCCTCG CTATGCCAGACAGCGAGAAATACCACTGAGCCCAACCCAGTGGTCACCCAACTCC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGAACTCACTGACGAGGAAAGGATGCTCTGGCATGATTCTCCAGATGAAGAATTCA CAACATACCTGGTGAAGAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTGA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGTTCGAAACTAAGGCTTCIGGTCTCCAGGAAGGTCAACAACTGCTGACTCTGGAC TGGTCCCGGCCACTAAAGCTTTGAACCTTCAGCAAGGACAGTGACTGCTGACTCCAGAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTAAATCATGGTTTTATGAAGAGTATAGCTTCTGGCAGAGTCTCAACTGTGAATC CAATATCTGATAGTCTATGACAAACATAAATTATGGCAGAACTCTCCGTTGTGAATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CACAAACACTACATCTGAGCACCTTCCTGATTCAAAATATTTAGTGTAAACAAAGGTA CACAAATACTACCACTGAGAAACCAAGATTCAAACCTGTCTACTGCTATTAAACTA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACAAACACC CAGCCATGTCAAAGAATATAAACCCAAACATGTCAAGCCAAATACAAGGCACAAACATC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAACCCCAATCATTATCTTTCCATC-----AGTAGCTGAATTCGAGATTCTG-CT AACATTATCACTGTCTTTCCACTGCTACTTGGAGCACTGAATTTAGGACTCTGACA -----

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGCA---GGAAGAGCATCTTCCCAAAGTGCAACCCCTGTACAGGGGGAAACATGGCT
GAGGGAAAGAGGAAGAGAGCATTTCC--AGTA-AOCCCCAATAACAGTAAGGACTATGATC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACCTATGGCCATACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTTC
AAAGATGNTCAATGTCAANATGCTTAGTAGCAACCAACAAA-CTATTA-----TTAG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AGCCATATATCCAAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCCAGCAGC
AGTCAGTAATACCAAAATAGTCAT-----CAGACATCTGTAGAGAGTGAAGTGAAC

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CTAGCAGTAGTGACATCTCTTCTCAGCTACTGCAGACCCCTAGCTTCTCCAGTCACCCCT
CCAGGCACATCACTTCTATTCTCAGCTACTCAAATACTTAGCACCTCCAGTTCCTT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGGTTACACACACCACTGCTCTCTTTATTTACATTCCTAGAAACAAACATACAGGTA
CAGATCCACACACAGCTGCTCATTTCTCAGTTTCGATCCCTAGANNNAATAGTACAGTTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

ACTTCCCTTGTCCAGGCCTTGGGAAGAGAGAGGACAATTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAGACGCTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGGGATT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

AAAACCCACATAGAACCCAGTTCTCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTCAGGTCAACAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TCAGGGACCTGCTAACAAAATGTGAGCCAGTTCCAGCCACAGAGTACCCCTGGGATGT
CCAGAGGTTCTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGTCAATGTGACAT

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GCCACACATGTCTTCCGAGAGGGGCTCAGTGGCTACTGCAGCACTGTCACTTCCAA
GTCTGTCTGTCTTCCAGGGAGAGGCTCACTGCTGCCACAGCAGCATTGTCTTTCCAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTTCTATCCACAGTGGCTCCCCAAACTAAATATGTGGGGTCTATGCAGAGAGTCTA
GTGCTGTCTCCATCACCTTCCCCAAGCTGACATTGCTAGAGTCCCATCAGAGAGTCTA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CCACTGTGGTCAAGAAACCACTGTACTATTTAAGGACAAACAAAATGTAGATATTGAGA
CAACTCTAGTCCAGATCCACTATTACTTGTAGAACAAACCCAGTGTAGANNNGAAA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

TAATACAAACCACTAATAAATATTCGGAGGGGAAGTAACCAAGTGATTCTACGGAG
NNACACACCCACAATAAATATTCAGGACTNGAAATTCOAAGTGAATCCAACTGGTG

rat_cDNA
human_5+3 corrected
mus_cDNA_5

CAGCATGACTTCTGCTCCACATCTGTATCCCTGGGGAATCTCTGTAGACAAATAGTG
CAGTCATGACATATGCTCCACATCCATACCCATGGAAAAAAGTCAAAAGTAAAGGCA

rat_cDNA
human_5+3 corrected
mus_cDNA_5

GTCACCTGAGCATGCTGGGACCATCCAACTGGGAAGATTCAAGTGAACAAACACCAC
GTTACCCACGTGTGTAGACCAATGAAGCTAAAGAGATTCAAGTGAATACATCGTCAC

0099430.030402

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTCCGAGCCCCCTCAGCACCCCTCAATACCA-----CAAGCACAAATTCTCAAGA
TTTCAGGTGCTATCACCAAGCCCAATGACTATTATAGCCATTACAGGTTTCAAGAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGAAACTCCCTTGCCACCAGATCTTTGTAATAACCAAGAAGSAGGGGATGTTAAAGA
GGAAATTCCTGGCAGACAGACTTTGTAATAACCAAGGCGGATTAAGGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGGTTCCCAAAATAGCTCCTC
ATCAACATAAGTTAGTTTACAAAAAGCACAGCTGTGATGCTTCTCTAAACATCTCCTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCGCGCA
CTTTACC-ACAGAGACAAAGTTCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAGGGCACTGAAGTAGTATCAGGT
CAATTCCTATCTATACCTTGACTACCGCTCACCACACTACGACAAA--ACACACAA-T

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCCAGAAGTCTCTCAGCAGGGGAAGAGCCGCTTCAACAACTCTCTCCAG-TGCTTCC
CCTGGAAGTCTTCCACAAAGAGGAGCTTCCCTTC-CCACCCCTTACCCCTATGCTTCC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TAGCACATAAGCAGAGATCTATACATTAAACTTCTTGTCAACGGAAACCCACAGT
TAGTATTATAAGCAAGACTCAAGTACAAAAGCATCATATCAACGCAACAGCAACCGC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACAAGTCCTACTGCTACTGCATCTGTCTATTGTCTGAAACCAACGACAGATCCAA
AACAACTCCTACCTTCCCTGCATCTGTCTATCTATGAAACCAACAGAGAGATCTAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGARGCAAAGACCAATTAAGGGGCTCG--GAGAACAGAAACAGGCAACACCCAC
AGCACAAACAATACAAAGAGAAGGAOCTCAAAGAGAACAGGACTGACCCAAACATCTC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCCCAGGCAGGTTTCTGGCTATAGTGCATCTCAGCTCTAACACAGCTGATACCCCTT
TCCAGACCAAGATTCTGGCTTCACTACCCCACTGCTATGACNACCTCCTNNGCTCTNN

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGCTTTCAGTCATTCCCCACGACAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA
NGCATTCACTCATTCGCCACAGAAAACACAACTGGGATTTCAAGCACAATCAGTTTTC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCAACAACCTCTCTCTGGCCA--TAACTGAAGTTTGAG-AGTAC--AACCAGAC
TTCAAGAACTCTTAATCTGACAGATGTGATTGAGAAGTACGCCAAGCAAGTACTCAGAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTGGGAATACAAAGCTTTGGAAACAAAGTTGTGAGCAATCAGAGAGATACCAC
TTTGAAGAGCACATTTGCTTCTGAAACAACTTTGTCCAGCAATCACCAGAGTACCAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGTGAAGAGAGCTCA--GACAC--ACCAACACCTCTCTCAGAGTGGGGGGCCCCC
AACTAGGAAGCATCATTAGACCTCAACCAACCACTTCTTGAGCAGCAGTGTCTACT

0001630-03400

TTGGAAGTACTGTGGAACTAAGTGCAGAGTGGAGGGTATGCCGAGGCCCTACGGTTTCT
CCGGAAGCACTGTGGAACTGAAGTGCAGAGCCAGAGGTAGGCCAAGCCCTACAGTTA

GCATACCTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAGGGAAGCAGAAAGGTCTGGG
GGATTCTTGCAAACCAAACAGTTGTCTCAGAAATCATCCAGGGAAGTAGGCAGGCTGTGG

TAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA
TGAOGGTTGACGGAACATTGGTCCTCCACATCTCAGTATTTATGACCGTGGCTTTTACA

AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACCTGTTGGTTAAGATACAAGTCATCA
AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACCTGCTGGTTAAATACAAGTCATTG

CAGCTCCCOCTGTCAATTATAGAGCAAAGAGGCCAAGCCATCGTTGGGGTTTATAGGTGCAA
CAGCACCACTGTATTCTAGAGCAAAGGAGGCCAGTCAATTGTAGGCACTTGGGGTGAAA

GTTTGAAACTGCOCTGCACTGCAAAGGAAGTCCCAGCCAGTGTTCAGTGGGTCTTT
 GTTTAAAGACTGCOCTGTACTGCAAAGGAAGTCTCCAGCCAGCGTTTACTGGGTCTCT

ATGATGGGACTGAACATAAAACCAATTGCAGTTGACTCATTCCAGATTTTCTTGTATCCAA
CTGATGGCACTGAAGTGAAACCAATTACAGTTTACCAATTCCAGTTGTTCTTAATTTTCAA

ATGGAACTCTGTATATAAGAAAGCATGCGTCTTTCAGTCAGGGGCACCTATGAGTGCATTG
ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACCTATGAATGCATTG

CCACACGCTCCTCAGGCTCAGAGAGAGGGTAGTGATTCTTACTGTGGAAGAGGGACAGA
CTACCAAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA

CCAATCCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT
 CCAGCOCAGGATAGAGCTGCATCCCAGAAAGGACTGAAGTGAATTTGGGGACAAT

TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT
TACTACTGAACTGCTCAGCCACTGGGGAGGCCAAACCCCAATAATGTGGAGGTTAOCAT

CCGAGGCTGTTCATCGACCAAGTGGCAGAGATGGGCGAGCCGAATCCAGGCTACCCAAATG
CCAGGCTGTGGTCCGACCAAGTGG-----GCGGCTGGATCCAGGCTACCCAAATG

GATCCTTGGTGGTTGGGTCAAGTGAACGGAAGAAGACGCTGGTGACTACTTATGTGTGGCAA
GATCCCTGTTTATTGGATCAGTAACAGAAAGAAGACAGTGGTGTCTACTTGTGTGTGGCAA

GAAACAAATGGGAGATGACCTAGTCCTGATGCATGTCCGCCGAGATTGACACCTGCCA
GAAACAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAAACCTGCCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAATTGAACAGAAGCAGTATTTTAAAGCAAGTGTCCATGGGAAAGATTTCAGTTG
AAATTGACCAACAGCAGTATTTTAAAGCAAGTGTCCATGGGAAAGATTTCAGTTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG
ATTGCAAGCTTCCGGCTCCCTGAGTGCAGAGATATCTGGAGTTTGCTGATGGAAOCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT
TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TCCACAATGGAACTTGTATTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATCT
TCAACAATGGAACTTTATCTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTCTGCCAGAACACCTTAGGGAAAGATGAGTGAATGTCACCTAACAGTTCTAACAG
GCTATGCCAGAACACCTTAGGGAAAGATGAATGAAGTCCACTTAACAGTTATAACAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCATCCCAAGGATAAGGCCAAGCTACAGACCAACATGAGGCTCAGGGCTGGGAAACAG
CTGCTCCCGGATAAGGCCAGTAAACAAACCAACAGAGAAATCAAGCTGGAGACACAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGTCTTGACTGCGAGGTCCTGCGGAAACGAGCCCAATGATTTTGGTTGCTGCTT
CTGTCTTGACTGCGAGGTCCTGCGGATCCCAACCAAAATATTTTGGTTGCTGCTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCAACAATGTCATTTCTTCTCCATGACAGGTTCAATTTCTATGCCAATAGACTTTGT
CCAATGACATGATTTCTTCTCCATGATAGGTACACATTTCTATGCCAATGGGTCTTTGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CCATCCATAAGTGAARCACTTGAATCTGGGACTATGTGTGCTAGCTCAGATCCTA
CCATCAACAAGTGAARCTGCTGATTTCTGGAGGTACGTATGTGTAGCCCAATCCCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GTGGGATGACACTAAGACATACCAACTGACATTTGTCTTAARCTCCATTATCAATG
GTGGGATGACACCAAAATGTACCAACTGATGTGTCTTAARCTCCATTATCAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCCTGTATGCAACAAGACTGTTATTAAGCCACAGCCATTCCGGCACTCCAAAAAATCT
GTCTGTATACCAACAGAACTGTTATTAAGCCACAGCTGTGAGACATTCCAAAAAATCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTGACTGCAGAGCAGATGGGATCCATCTTCCAGGTCACTGGATTATGCCAGGCAATA
TTGACTGCAGAGCTGAAGGGACACCATCTCTGAAGTCATGTGGATCATGCCAGGCAATA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TTTTCCTCCAGCTCCATCTTTTGGAGCAGATCAAGTCCATCCAAATGGAACCTTGG
TTTTCCTCAGGCCCACTATGGAAGCAGAAATCAGATCCATAAATGGAACCTTGG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGATGAGGAACATCCGCTTTCTGACTCTGCGGACTTCACTGTGTGGTTCGGAGGAGG
AAATTAGGAATGTGAGGCTTTCAAGTTCAGCCGACTTTATCTGTGTGGCCCAATGAG

0901630-030402

rat_cDNA
human_5+3_corrected
mus_cDNA_5
GAGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGGAAGACCGACAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACT
TTAGAAATCCATTTAATGAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
GCTCTGTGGATGGGAACCCCCACCTGAAATTAOCTGGATCTTACCTGAOGGCACACAGT
GCTCTGTGTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAATGGCACACGAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TTGCTAACAGAACACAAATTCOCCTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
TTTCCATGGACCAAAAGTTATCAGTATCTGATAGCAAGCAATGGTCTTTTATCATT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAACAACCTCGGGAGGATGCAGGAAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACATCGAGAACTCATCTGTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5
CAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTCTCTGATGGGA
CAGGGACAGTAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTCTGATGGAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATGACAGGCOCTCAAG
TCCCTAAGCCAAATATCAAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCOCTCAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TGGATGGAAATACATACTGCATGAAATGGCAGCTGGTCATCAAGCAACACAGCTC
TTAATGGGAATACATATTGCATGACATGGCAOCTTAGTCATTAAAGAGCAACAGCTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACGACCAAGGAAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCACTTATTAGCG
ATGACAGAGCAAACTATATCTGTAGGCTCAAAATAGTGTGGTCATACACTGATTACG

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TGTCAGTGATGGTGTGTGGCTACCOCTCCCGAATCATAACTAOCCTACCCAGGAACATGC
TTCCAGTAATGATTGTAGCOCTACCOCTCCCGAATTACAAATCGTCCACCCAGGAGTATTG

rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCCAAGCCAA
TCACCAAGGACAGGGGCAGCCTTTGAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5
AAGTCACCTGGGAGACGCGAAGACACTCOCTGCTCTCAAAAGCAACAGCAAGAAAAACCC
AAATCACATGGGAGATGCCTGACCACTCOCTTCTCTCAACGGCAAGTAAGAGAGGACAC

rat_cDNA
human_5+3_corrected
mus_cDNA_5
ATAGAAGTGAGATGCTTCACCCCAAGGTACGCTGGTCATTGAAATCTCCAAACCTCGG
ATGGAAGTGAGCAGCTTCACTTACAGGTACCOCTAGTCATTGAAATCCCAACCTCCG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATTCCGGAGTCTATAAGTCAGAGCTCAGAACCTACTTGGGACTGATTACGCCAACACTT
ATTCTGGGATATACAAATGCACAGCAAGAACCCTTGGTAGTGATTATGCAGCAACGT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ACATCCAGGTAAGTCTGACAGGAGGGGGAGACTAAAATTCACAGAGTCCACATCCCA
ATATTCAAGTAATCTGACATGAA-----ATAATAAAGT-CAACAA-----CATCTGGGCA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGGTTTATTTTTTGAAGAAGTTTAAATCAAGGCGCCATAGGCATGTAAATGAGTCTGA
GAATTATTTTTTGAAGAAGTTTAAATCAAGGCGCCATAGGCATGTAAATGAATTGA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATACATTTCAGTATTAAATTTACAATGGACATGGA--TGA---GACTTGTAAATGAAA
ATACATTTCAGTATTAAATTTACAATGACATGCAAAATAAAGGACTTGTAAATTAAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCATTGTGAAGTGA--ACCGAGTCTCTG--TGGATCTCAAAGCAAACCTCTTAACCTAA
GCATTATGAAGTGAATGACTGATTTATTTAATGATCTCAAAACAACTTTTAACCTAA

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GGCAGTTTGATTTTGCCAAACAAATATAACAAACATTAAGAGAAAAAATGATCCACTAC
GGCAGTTTTATTTTGCCAAACAAATAACAATAACAA--ACATTGAAACGGTTCACTAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GAAATACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTTCTCTCGCTAAC
AAAATAACAAATGGCTAATGTACCTGAATTTTCAGTAAAAAA--TGAAGTT-CTAAT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGTTGCCAGCTGCCTCGTGTCTGTTTCCCTACCAATGTCAACACATCCGACACAGGGTGA
A---CCAGTTGCCATAGTGTCCACCTCTCATCAATGTTACAAGCATGGCACTCAG---A

rat_cDNA
human_5+3_corrected
mus_cDNA_5

ATGGAGTCAACGGGAAGATTAAAGTTTCCGGTCTGTGTAATCTCAATGTACAAATATTC
ACAGAGACATGGAAAAATATTAATCTGC-----AATCT--ATGTATAAATATTT

rat_cDNA
human_5+3_corrected
mus_cDNA_5

TGTCNCTGGTTTATAAACATTTT-GATAAAACCGAAAAAATAAAAAAAAAAAAAAAAAA
TGT----GGTTTATAAATTTTTTGTAAACCTACAGAAATAAG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAA

(rat_cDNA: SEQ ID NO:7)
(human_5+3_corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15

rat	MQVGRGVSGLLISLTAVCLVVPFSGSRACPRRCACYVTEVHCTFRYLTSIPDGIPANVE
human_5+3_corrected	MKVKGRTITCLLVSEFAVICLVATPGGKACPRRCACYMPTVHCTFRYLTSIPDSIPNVE
mouse_5_corrected	MOKRGREVSCLLISLTAVCLVVPFSGSRVCPRRRCACYVTEVHCTFRDLTSIPD-GPANVE
rat	RINLGYNSLTRLTENDFDGLSKLELIMLHSHNGIHRVSDKTFSGLSQLQVLKMSYNKVQII
human_5+3_corrected	RINLGYNSLVRLMETDFSGLTKELELIMLHSHNGIHTIPDKTFSDLOALQVLKMSYNKVRKL
mouse_5_corrected	RVNLGYNSLTRLTENDFDGLSKLELIMLHSHNGIHRVSDKTFSGLSQLQVLKMSYNKVQII
rat	RKDTFYGLGSLVRLHLDENNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTEVSLSYLQ
human_5+3_corrected	QKDTFYGLRLSLTRLHMDENNIEFINPEVYGLNRLVHLEGNRLTKLHPDTEVSLSYLQ
mouse_5_corrected	EKDTLYGLRLSLTRLHLDENNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTEVSLSYLQ
rat	IFKTSFIKYLELSDNFLTSLPKEMVSYMPNLESILYHGNPWTCDCHLKWLEWMOGNPDI
human_5+3_corrected	IFKISFIKLYLSINFLTSLPKEMVSYMPDLDSLYLHGNPWTCDCHLKWLESDWIQ--PDV
mouse_5_corrected	IFKTSFIKLYLYDNF-TSLPKEMVSSMPNLESILYHGNPWTCDCHLKWLEWMOGNP--
rat	IKCKKDRSSSSPQQCPLOMNPRIKGRPFAMVPSGAFCTKPTIDPSLKSKSLVTOEDNG
human_5+3_corrected	IKCKKDRSPSSAQOCPLOMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLFILEDSS
mouse_5_corrected	-----
rat	SASTSPQDFIEPFGSLSNMTXSGNKADNVCSIQKPSRTSPATTEENDYTMINASFTST
human_5+3_corrected	SAFISPGGFMAPFGSLTINMTDQSGNEANMVCSIQKPSRTSPIATTEENDYIVLNTSFTST
mouse_5_corrected	-----
rat	NIVCSVDYNEHQPVWQLALYSDSPLILERKPKLTETPSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHQIPVWQILALYSDSPLILERSHLLSETPOLYYKYKQVAPKPEDIFTNIEA
mouse_5_corrected	-----
rat	DVRADFFWFQOEKIVLQNRATATLSTLQIQFSTDAQIALPRAEMRAERLKWTHILMMNN
human_5+3_corrected	DLRADPSWLMQDQISLQNRATATFSTLQIQYSSDAQITLPRAMRPVKHKWTMISRDNN
mouse_5_corrected	-----
rat	PKLERTVLVGGTIALSCPGKGDPSPHLENLADGSKVRAPYVSEDGRILIDKNGKLELQM
human_5+3_corrected	TKLEHTVLVGGTVGLNCPGQGDPTFHVWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
mouse_5_corrected	-----
rat	ADSPDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDELPCISTGV
human_5+3_corrected	ADSDTGVYHCISSNYDDADILTYRITVVEPLVEAYOENGIHHTVFIGETLDELPCSTGI
mouse_5_corrected	-----
rat	PLASISWILPGNTVFSQPSRDRQILNNGTLRLQLQVTFKQGGHYQCVANPSCADFESEFKV
human_5+3_corrected	PLASISWVIPGNVLYQSSRDKKVLNNGTLRLQLQVTFKQGGYTRCVANPSCGVDFLIQV
mouse_5_corrected	-----
rat	SVQKKGQRNVEHREAGGSGLGEPMSSVSLKQPSILKLSASALTGSEAGQVSGVHRKNK
human_5+3_corrected	SVKMKGRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGHKTSSTSKRNN
mouse_5_corrected	-----
rat	HRDLIHRRRGDSILRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPPKQENTTVKVP
human_5+3_corrected	YRELTLQRRGDSILRRFREHRRQLPLSARRIDPQWAALEKAKKNNAMPDKRENTTVSFP
mouse_5_corrected	-----

20190923 16:33:33

0001630 030402

rat	FLAVPLVELTDEEKDASGMIPPEEFMVLTKASGVPGRSPTADSGPVNHCMTSIAST
human_5+3_corrected	PVVTOLNIPGEEDDSSGMLALHEEFMVPATKAINLPARTVTADSRITSDSMNTNINYGT
mouse_5_corrected	-----
rat	EVS-TVNPQTLOSEHLPDFKLFVYNGTAVTKSMNPSLASKIEDTTNQNPIILFP---SV
human_5+3_corrected	EFSPVYNSQILPPEEPTDFKLSTAIKTAMSKNINPTMSSQIQGITNQSSTVFILLGA
mouse_5_corrected	-----
rat	AEIRDSAQAGRAS--SQSAHPVTGGMATYGHNTYSSFTSKASTVLQPINPTESYGFQI
human_5+3_corrected	TEFQDSQMGGRGHEHFQSRPPITVRIMIKOVNVKMLSSTTNKL--LLESVNTTNSH--QT
mouse_5_corrected	-----
rat	PITGVSRPSSSDISSHTADPSTFSSHPSGSETTASLETHIPNNNTGNFPLSRHLGRERT
human_5+3_corrected	SVREVSEPRENHFYSHTTQILSTSTFSPDPTAHSQFPIPRNS-TVNIPLFRFRGRQRK
mouse_5_corrected	-----
rat	IWSRGVKNPHRTFVLRHRHRTVREPAIKGPANKNVSVQVPATEYPMCHTCPSAEGLTVA
human_5+3_corrected	IGGRGRIISPYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRERLTTA
mouse_5_corrected	-----
rat	TAALSVPSSSHSALPKTNVGVIAEESTTVVKKPLLLFKDKONVDIEIITTTTKYSGGES
human_5+3_corrected	TAALSFPSAAPITFPKADLARVPSEESTTLVQNPLLLLENKP--SVEKTPITIKYFRTET
mouse_5_corrected	-----
rat	NHVIPTASMTSAPTSVSLGKSPVDNSGHLMPGTIQTGKDSVETTPSPSLSTP---SIP
human_5+3_corrected	SQVTPTGAVMTYAPTISIMEKTHKVNASYPRVSSSTNEAKRDSVITSSLSGAIKTPMTII
mouse_5_corrected	-----
rat	ISTKFSKRKTPLEHQIFVNMOKKEGMLKNFYQFGLQKNFAAKLFXIAPLLPTGQSSPSDST
human_5+3_corrected	AITRFSRKKIPWQONFVNMHNPKGFLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFT
mouse_5_corrected	-----
rat	TLLTSPPPALSTMAATONKGTVEVVGARSLSAGKKQ-PFTNSSPVLSPSTIEKRSNTLNF
human_5+3_corrected	TLSTSVMOIPSNLTATAHTTTKTHNPG-SLPT-KKELPPFPLNFMPLPSIISKDSSTKSI
mouse_5_corrected	-----
rat	ISTETPT-VTSPTATASVIMSETQRTSRKEAKDQIKG-P-RKNRNNTTTPROVSGYSAY
human_5+3_corrected	ISTQTALPATTPTFPASVITYETQTERSRQAQTIQEQEPQKKNRTDPNISPDQSSGFTTF
mouse_5_corrected	-----
rat	SALITADTFLAFSHSPRODDGCVSAVAYHSTTS--LJAITELFEKTYTTLGNTTALETT
human_5+3_corrected	TAMTP-PALATHSPPENTTGISSTISFHSRTIINLTDVIEELAQASTQTLKSTIASETT
mouse_5_corrected	-----
rat	LLSKSQESTTVKRAS-DTP-PFLLSSGAPFVPTSPFPPTKGVVTDGKVTSFAQMTSNRV
human_5+3_corrected	LSSKSHQSTTTKASLDTPIPPFLLSSSATIMPVPISPFPTQRAVTDTRGDSEFRIMINTV
mouse_5_corrected	-----
rat	VTIYESSRHNTDLQQPQAEASPNPEIITGTTDSPNLFPSTSVBALRVDPKONSKEKESP
human_5+3_corrected	VKLHSSRHN--LQMPSSQLEP-----LTSSTSNLLHSTPMPLATTVKSONSKLATSP
mouse_5_corrected	-----
rat	WPEIKYOLKSYSETIEKGRPAVSMSPHLSPLEASTHASEWNTQKHAESVFDKPKGQNP
human_5+3_corrected	WAEYQFWHKPYSDIAEKKGKPEVSMLATTLGLSEATTLVSLWDGCKNTKSDFKKPVQEA
mouse_5_corrected	-----

Figure 16

rat	MQVRGREVSGLLISLTAVCLVVTFGSRACFRRACACVYVTEVHCTFRYLTSIPDGPANVE
human_5+3_corrected	MKVKGRTGTCCLLVSTFAVICLVATPGGKACFRRACACVYVTEVHCTFRYLTSIPDSIPENVE *:*** :: **:*::::***.***:*****;*****;*****;***.***
rat	RINLGYNSLTRLTENDFDGLSKLELIMLHSENGIHRVSDKTFSGLSLOVLKMSYNKVQII
human_5+3_corrected	RINLGYNSLVRIMETDFDGLSKLELIMLHSENGIHTIPDKTFSDLOALQVLKMSYNKVRL *****.***.***.***:*****;*****;*****;*****;*****: :
rat	RKDTFYGLGSLVRLHLDENNIEFINFEAFYGLTSLRLVHLEGNRLTKLHPDTFVLSYIQ
human_5+3_corrected	QKDTFYGLRSLTRLHMDENNIEFINFEVYGLNRLVHLEGNRLTKLHPDTFVLSYIQ :***** **.****:*****;****.*****;*****;*****
rat	IFKTSFIKYLFLSDNFLTSLPKENVSYMPNLESILYLGNPWTCCHLKWLSEWMOGNPDI
human_5+3_corrected	IFKISFIKELYLSNFLTSLPKENVSYMPDLDSLYLGNPWTCCHLKWLSDWIQ--PDV *** ***:***:*****;*****;*****;*****;*****;*****:***
rat	IKCKKDRSSSSPQQCPCLMNPRIKGRPFAMVPSGAELCTKPTIDPSLKSLSLVTQEDNG
human_5+3_corrected	IKCKKDRSPSSAQCPCLMNPRTSKKPLAMVSAAPQCAKPTIDSSLKSKSLTILEDSS *****.***.***** ***:***.***.***.***.***.***.***.***.***.***
rat	SASTSPQDFIEPFGSLSNMTXSGNKADMVCSIQKPSRTSPTAFTRENDYIMLNASTFT
human_5+3_corrected	SAFISPGGFMAPFGSLTLMNTDQSGNEANMVCSIQKPSRTSPTAFTRENDYIVLNTSFTT ** ***:***:***** ***:***:*****;*****;*****;*****;*****
rat	NLVCSVDYNIHQPVWQALLALYSDSLILERKPOLTEPSSLSSRYKQVALRPEDIFTSIEA
human_5+3_corrected	FLVCNIDYGHIQPVWQILALYSDSLILERSHLLSETPQLYYKYKQVAFKPEDIFTNIEA ***:***.*****;*****;*****. *:***.***.***.***.***.***.***
rat	DVRADPFVFOQEKIVLQNLRTATLSTLQIQFSTDAQIALPRAEMRAERLKWMTIMMNN
human_5+3_corrected	DLRADPSWLMQDQISLQNLRTATLSTLQIQYSSDAQITLPRAMRFVKHWTMTISDRAN *:*** *:***.***.***.***.***.***.***.***.***.***.***.***.***
rat	PKLERTVLVGTTIALSCPGKGDPSHLEWLLADGSKVRAPYVSEDGRILIDKNKLELQM
human_5+3_corrected	TKLEHTVLVGTTVLGNCFGQDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSKLELQM .**:***.***.***.***.***.***.***.***.***.***.***.***.***
rat	ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDPCLSTGV
human_5+3_corrected	ADSFDTGVYHCISSNYDADILTYRITVVEPLVEATQENGIRHTVFIGETLDPCHSTGI *****:***.***.***.***.***.***.***.***.***.***.***.***.***
rat	PDASISWILEPNTVFSQPSRDRQILNGLTLRLQVTFKQGHYQCVAAANPSGADFSFKV
human_5+3_corrected	PDASISWVIFGNVLYQSSRDKVLMNGLRLRLQVTFKQGYRCVAAANPSGVDELIFQV *****:***.***.***.***.***.***.***.***.***.***.***.***.***
rat	SVQKKQRMVEHREAGSGGLGEPNSSVSLKQFASLKLASALTGSEAGKQVSGVHRKNK
human_5+3_corrected	SVKMKQRPLEHGETEGSGGLDESPIALHKEPPGAQLRTSALMEAEVGKHTSSTSKRN **.****.***.***.***.***.***.***.***.***.***.***.***.***.***
rat	HRDLIHRRGDSTLRRFREHRRQLPLSARRIDPQWAAALLEKAKKNSVPKKQENTIVKPV
human_5+3_corrected	YRELTLQRGDSHRRFREHRRFPSPARRIDPQWAAALLEKAKKNAMPDKKENTIVSPP :*.***.***.***.***.***.***.***.***.***.***.***.***.***.***
rat	PLAVPLVELTDEEKDASGMIPPEEFMVLTKASGVPGRSPTADSGPVNHGEMTSTASGT
human_5+3_corrected	PVVTLPLNIPGEEDSSGMLALHEEPNVFATKALNLPARTVTADSRITSDSPMTNINXT *..* :..*.*.***.***.***.***.***.***.***.***.***.***.***.***
rat	EVS-TVNPQTLQSEHLPDFKLSVTNGTAVTKSMNPSIASKIEDTINQNPITIFP---SV
human_5+3_corrected	EFSPVNSQLRPEEPDFKLSSTAIKTAMSKNINPTMSSQIQGTTNQHSSTVFFLLGA *.***.***.***.***.***.***.***.***.***.***.***.***.***.***
rat	AEIRDSAQAGRAS--SQSAHPVTGGNMATYGHNTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected	TEFQDSQMGREGHEFPSPITVRIKDKVNVKLSSTTNKL---LLESVNTTNSH--QT *:***.***.***.***.***.***.***.***.***.***.***.***.***.***

20190103 03:05:55

rat	PITGVSRPSSSDISHTTADPSFSSHPGSGHTTASSLFIIPANNNTGNFPLSRHLGRERT
human_5+3_corrected	SVREVSEPRHNFYSHTTQILSTSTFSPDEHTAAHSQFPPIPNNS-TVNIPLFRFRGRQRK .: **.* ..: **** * *;.*..**.* * * ****. * *.* *;***.*.
rat	IWSRGRVKNPHRTFVLRHRHRTVRPAIKGPANKVNSQVPATEYPMCHTCPSAEGLTVA
human_5+3_corrected	IGGRGRIISPYRTFVLRHRHRTVRPAIKGPANKVNSQVPATEYPMCHTCPSAEGLTVA * .***: .*:*****; .*:.*:.*:..** * : * . * **.*
rat	TAALSVPSSSHSLPKTMNVGVIAEESTTVVKPLLLPKDKQNVDEIITTTTKYSGGES
human_5+3_corrected	TAALSVPSAPITFPKADIARVPSEESTTLVQNPILLLENKP--SVEKTPPTIKYFRTET *****,**.: :;***: . * ;*****;*;*****;:* :.* *.* ** *
rat	NHVIPTASMTSAPTSVSLGKSPVDNSGHLMPGTIQTGKDSVETTPPLSPLSTP--SIP
human_5+3_corrected	SQVTPTGAVHTYAPTSPINMEKTHKVNASYPRVSTNEAKRDSVITSSLSKATKPMITII .:* ** * ** ****;.: * : *!.: :..* : : *** *;.*.....* : *
rat	TSTKFSKRKTPHLQIFVNNQKKEGMLKNPYQFGLQKNPAKLPKLAFLPTGQSSPDSST
human_5+3_corrected	AITFSRRKIIFWQCNFVNNHNPGRRLRNOHKVSLQKSTAVMLPPTSPALFORQSSPHFT ; *;*.** * ;* ****;: * *.* :!..***..* . *** : * ** **** . *
rat	TLTSPFPALSTTMAATONKGTVEVSGARLSAGKQPTNSSPVLPTISKRSNTLNF
human_5+3_corrected	TLTSPVQIPSNLTHTTHTTKTHNP-SLPTKKELPFPPLNPLPSIISKDSSTKSI ** ** *.*:;!!; *.. . **.: *; **. .*:*** ** *.* :;:
rat	STETPT-VTSETATASVIMSETQRTSKEAKDQIKG-P-RKNRNANTTTPROVSGYSAYS
human_5+3_corrected	STOTAIPTATPTFPASVITYETQTERSRAQTIQREOEPOKQNRDTPNISFDQSSGFTPT **.*. .:*** .**** ** *;: . * : * ;***.: * * * *;: :
rat	ALTADTFLAFSESPRODGGNVSAVAYHSTTS--LHATELFEKYTOTLIGNTALETTL
human_5+3_corrected	AMTP--PALAFTHSPPEMTTGISSTISFHSRTLALTUVIEELAQASTQTLKSTIASETTL *.*. .:***:*** :; * *;:;*** * . * ** : **** . * * ****
rat	LSKSOESTTVKRAS-DTP-PPLSSGAPPVPTSPPPPTKGVVTDKSVTSAFQMTSNRVV
human_5+3_corrected	SSKSHQSTTRKASLDTPFPFLESSATLMPVPISPPPTQRAVTDTRGDSHFRLMTNTVV ***.:***;.* ** *;***.* . *.* .****: .***: * *;: ; * **
rat	TIYESSRHNTDLOQPSAEASNPETITGTTDSPSNLFPSTSVPALRVDEKQNSKNKPSFW
human_5+3_corrected	KLHSSRHN--LQMPSSQLEP-----LTSSTSNLHSTMPALTTVKSQNSKLPSPFW .:***** ** *;:.* *.*.***: **.*** . *.*** .****
rat	PEHKYQLKSYSETIEKGRPAVSMSPHLSLPEASTHSHWNTQKHAESVFDKPKQNP-
human_5+3_corrected	AEYQFWHKPYSDIAEKGRKPEVSMLATTLGLSEATLVSDDWGQNTKESDFDKPVQEAT .*;: *.*: ****.* ** , .*.**.*.*;: **;:*** **** *;:
rat	TSKHLPEVSLPKTILKKPRIIGGKAASFTVPANSDFLPCAVGDPLPIHNTVSSGXE
human_5+3_corrected	TSKLLPFDSLRSYIFEKPRIYGGKAASFTIPANSDAFLPCAVGNPLETHWTVS-GLD *** **; *;.: :;***;*****;*****.*****;*** ***** * :
rat	ISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNLGVDHFHVLSLVVFPARILDR
human_5+3_corrected	LSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCASNLFGTDHLVTLVSVVYPPRIER :*.*.*;:*****;***** * :.*.*;:*** **.***.*
rat	HVKEITVHFGSTVELKCRVEGMPRTVSNILANQTVVSETAKGRKVVVTPDGLTILYNL
human_5+3_corrected	RTKEITVHSGSTVELKCRAGRPSPPTVTWILANQTVVSESSQGRQAVVTVDGLVLHENL ;.****** *****.* * **;*****;:;***. ** *****:.*
rat	SLYDRGFYKCVASNPQGDSLLVKIQVITAPPVILEQKQAIIVGLGSLKLPCTAKGTP
human_5+3_corrected	SIYDRGFYKCVASNPQGDSLLVKIQVIAAPPVILEQKQVIVGTWGESLKLCTAKGTP *;*;*****;*****.*****;***.*. * *****
rat	QPSVHWVLYDGTTELKPLQLTHSRFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVV
human_5+3_corrected	QPSVHWVLYDGTTELKPLQLTHSRFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVV ***;*** ****;*;:.*;:*****;.*.* *****;*****
rat	ILTVEEGTIPRIETASQKNTEVNLGKILLNCSATGDPKPRIIWRLPKSAVIDQNHMRG
human_5+3_corrected	MLTNEERTVSPRIEASQKTEVNFQDKILLNCSATGEPKQIMWRLPKSAVIDQ----G ;*.** * ****;*** **;.*;*****;***.*;*****.* *

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDFTVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGNPWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTIPDSIPPNVE
RINLGYNLVRMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDFTVSLSYL
QIFKISFIKFLYLSDNFTSLPQEMSYPDLDSLLYLHGNPWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCISQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIQPVWQILALYSDSPILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEVRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTG VYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHTVFIGETLDLPCHSTGIP
DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYRCVAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVKGHTSSTSKRH
NYRELTQRRGDSTHRRFRENRRHFPPSARRIDPOHWAALLEKAKKNAMPDKRENTTV
SPPPVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
YGTEFSPVVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPTVTRTMIKDVNVKMLSSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSPHTAAHSQFPIPRNSTVNIPLFRFRGRQRKIGGRGRIIS
PYRTPVLRHRHYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT
AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETLSSKSHQSTTTTRKASLDTPIPPFLSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLMQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAF
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCASNL
FGTDHLHVTLVSVSYPPRILERRTKEITVHSGSTVELKCRAGEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLSIYDRGFYKCVASNPGGQDSSLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIEAASQKRTEVNFGDKLLLNC SATGEP
 KPQIMRLPSKAVVDQGSWIHPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTN
 KRIKAGDTAAVLDC EVTGD PKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVS KPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTP
 SPEVMWIMP DNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGE SVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSV DGNPPPEIIWILPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVGYIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHC VSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGN YICKA
 QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGA AFQLHCVALGV PKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

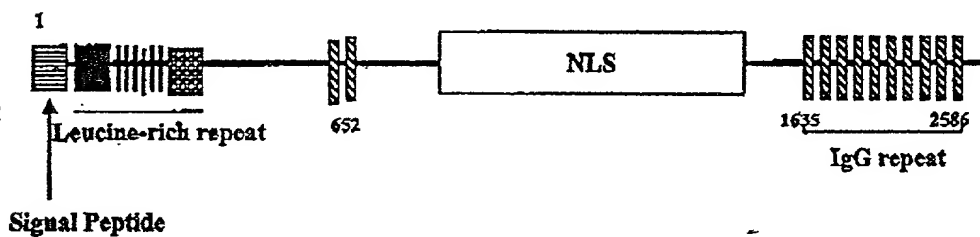


Figure 20

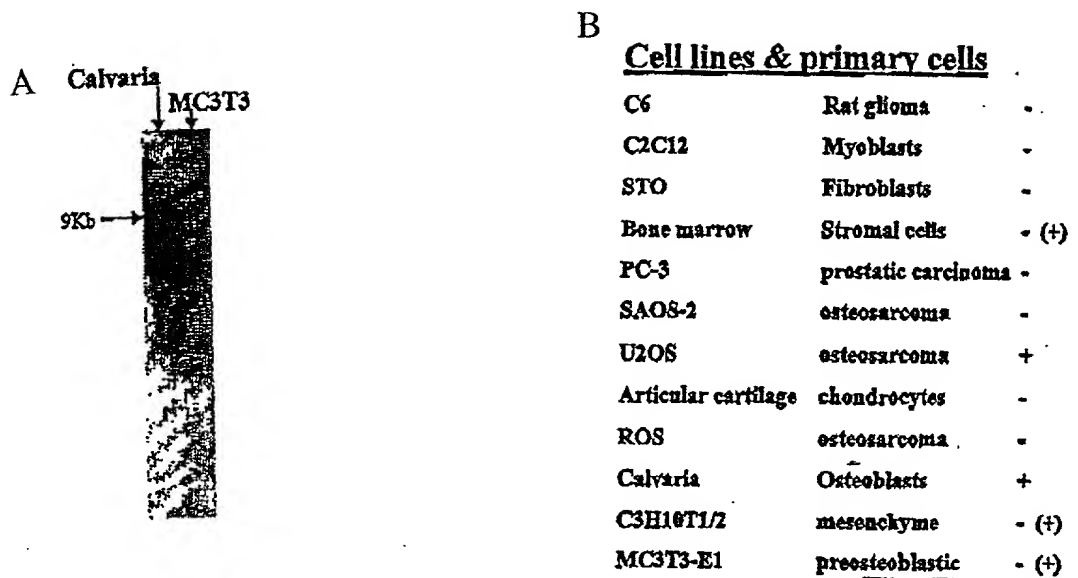


Figure 21

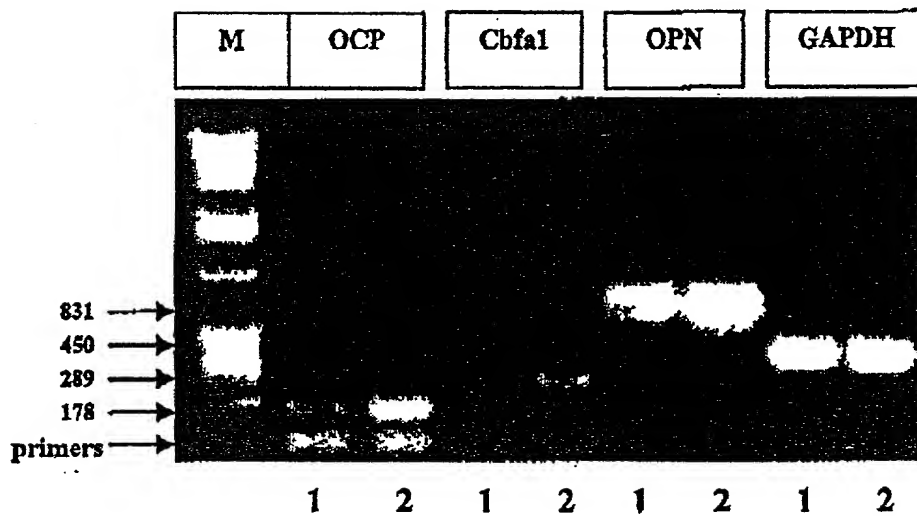


Figure 22

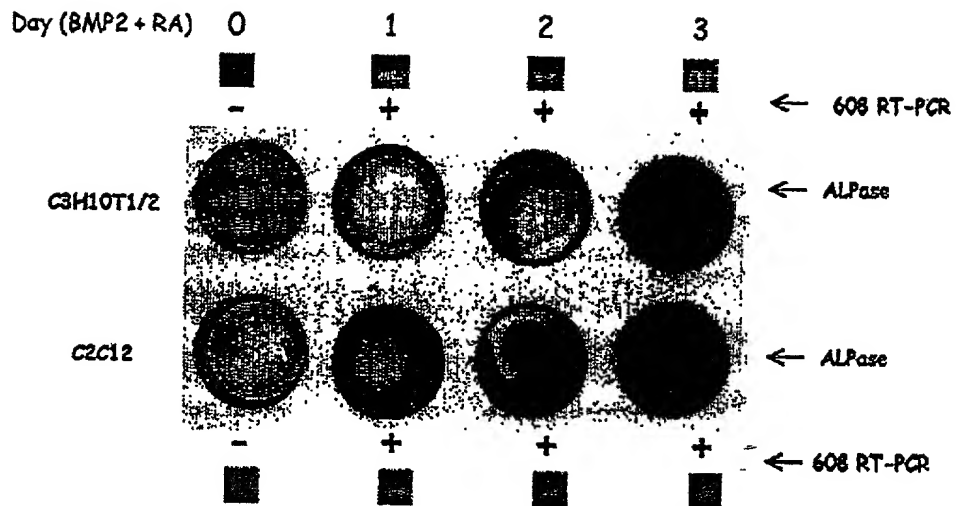


Figure 23

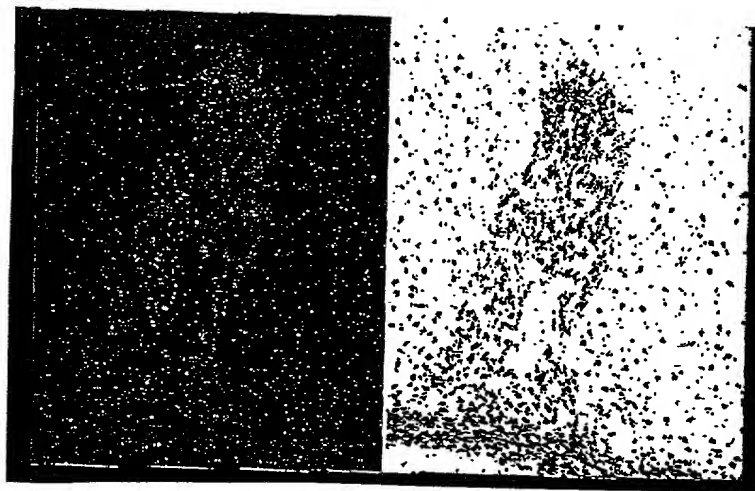


Figure 24

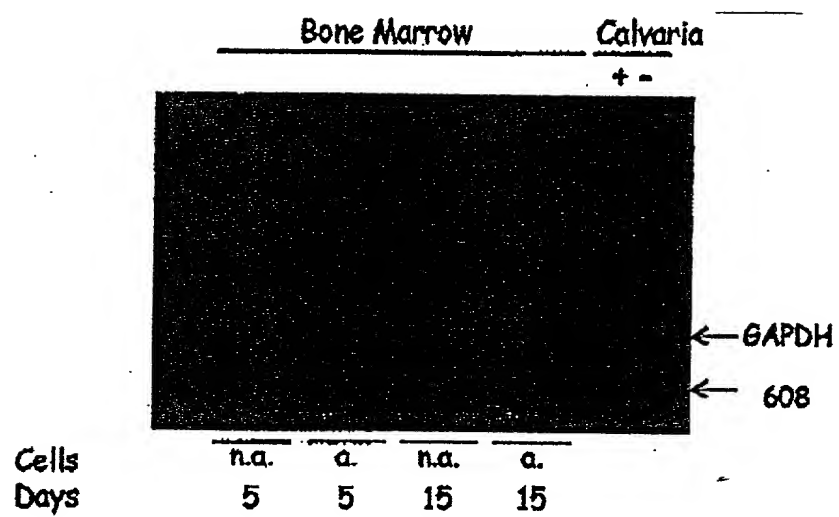


Figure 25

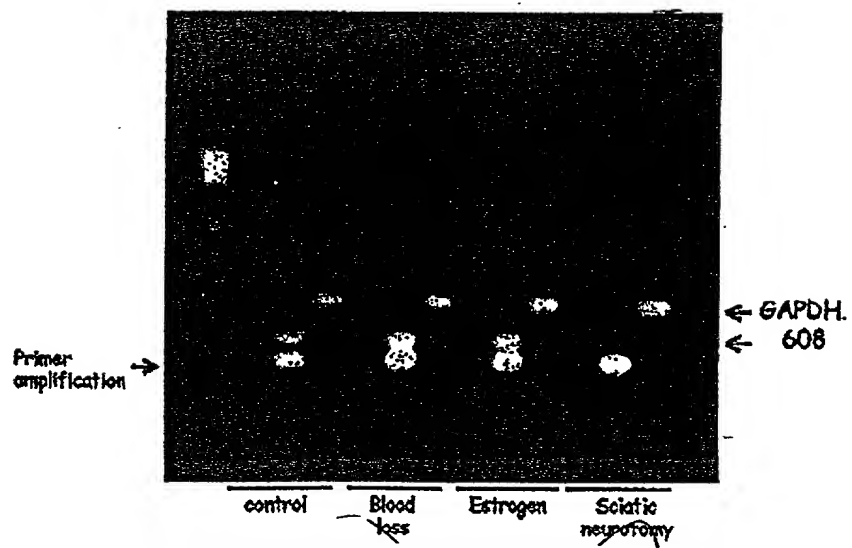


Figure 26



204030 00015562

Figure 27

A

B

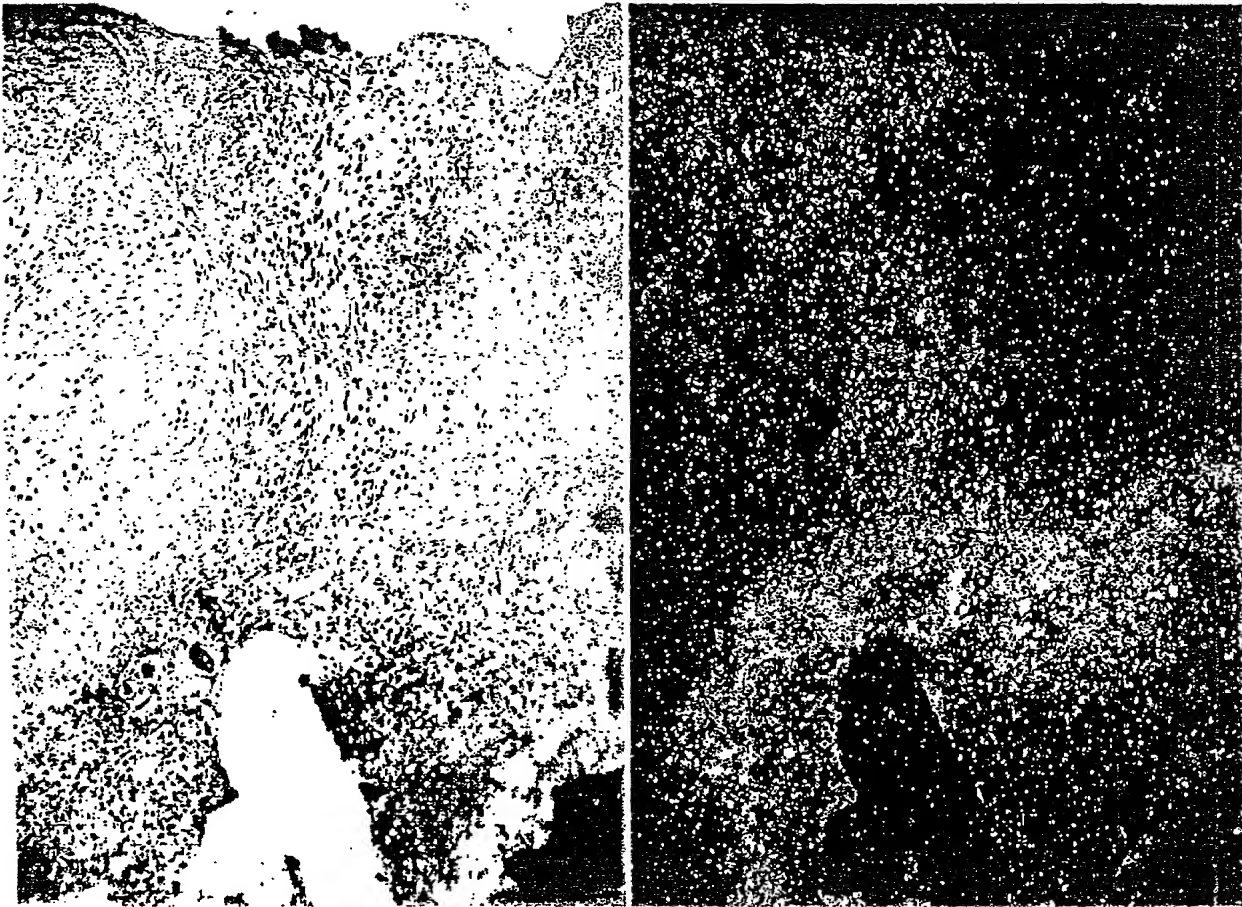


Figure 28

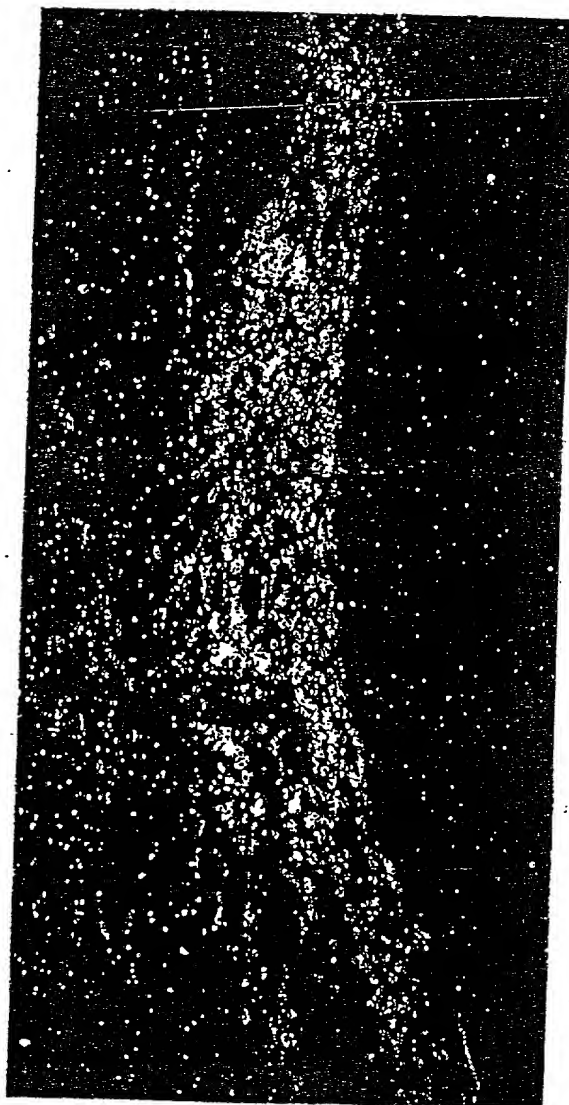


Figure 29

A

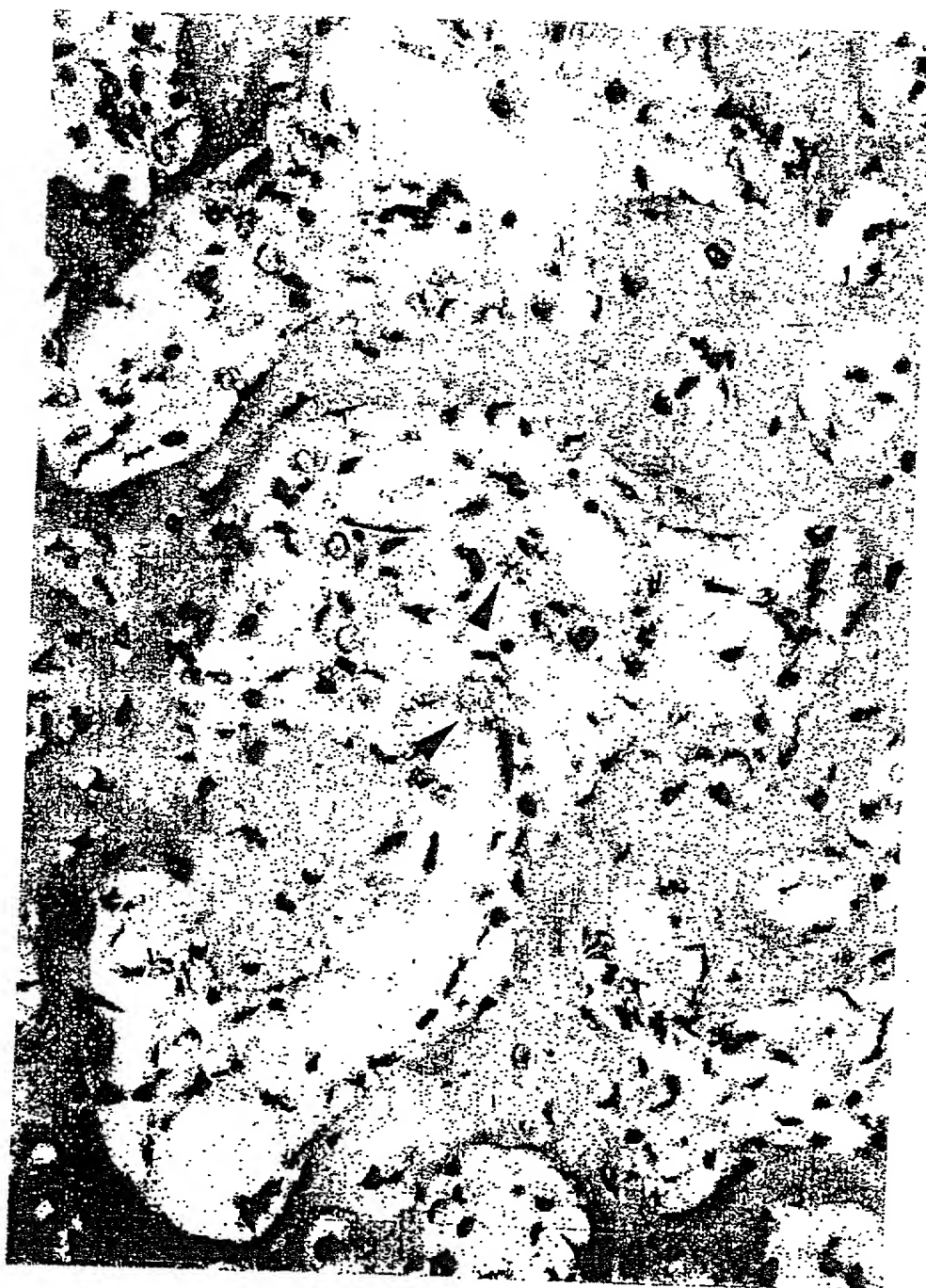


B



00991E30 020402

Figure 30



204060 0631660

Figure 31



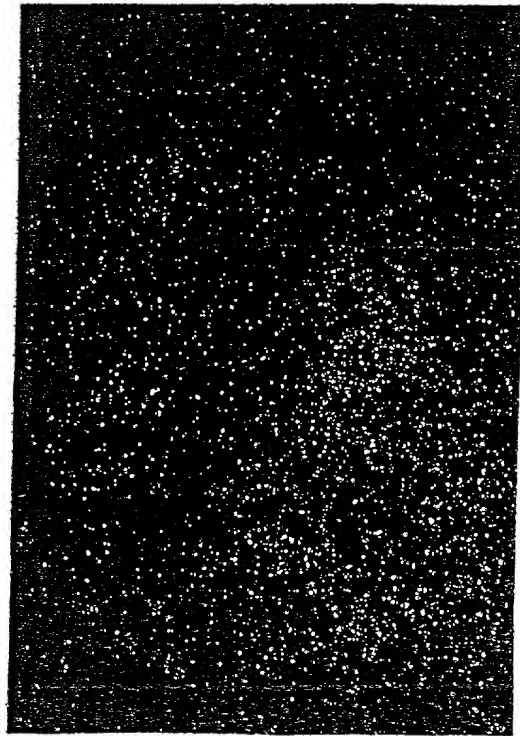
2040ED 0E9T660

Figure 32

A

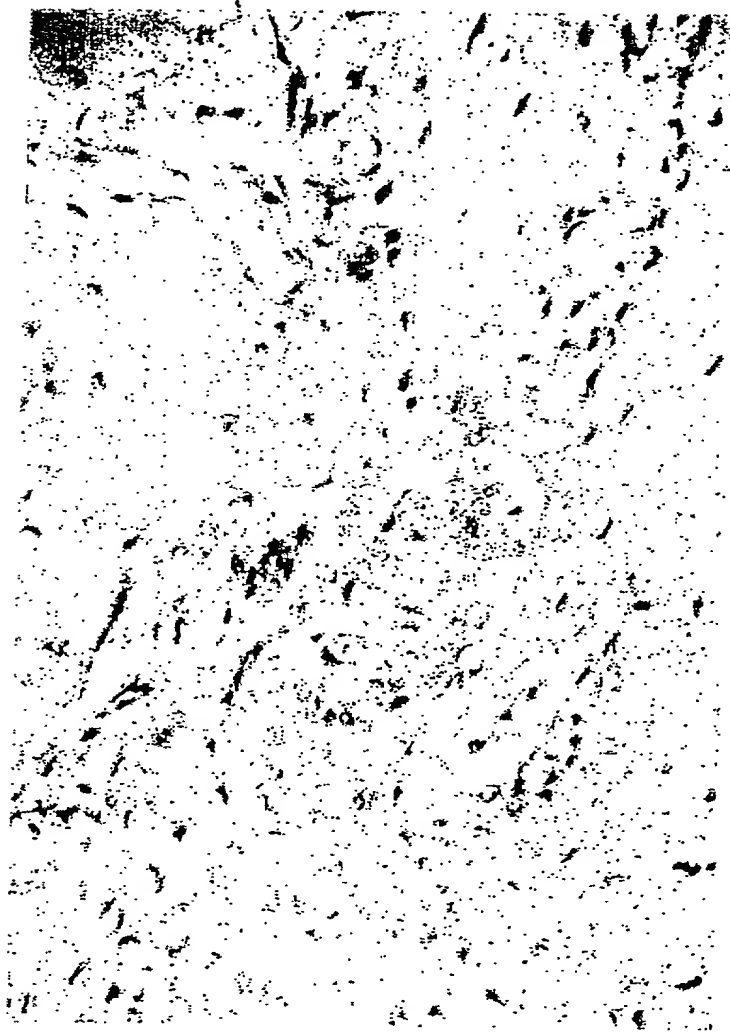


B



20200000 00000000

Figure 33



204000 00000000

Figure 34

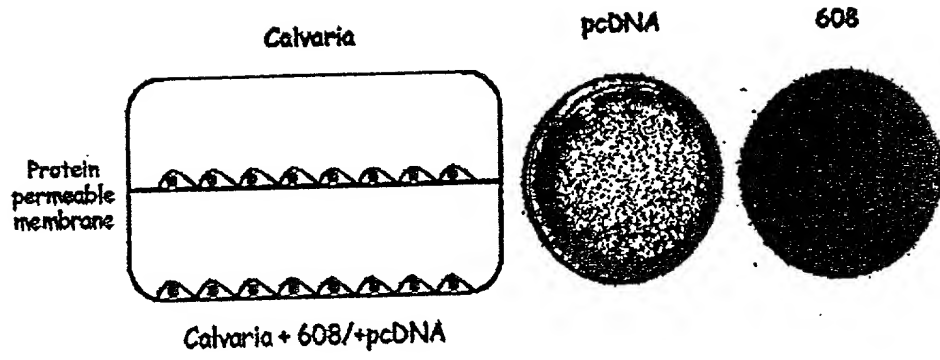


Figure 35

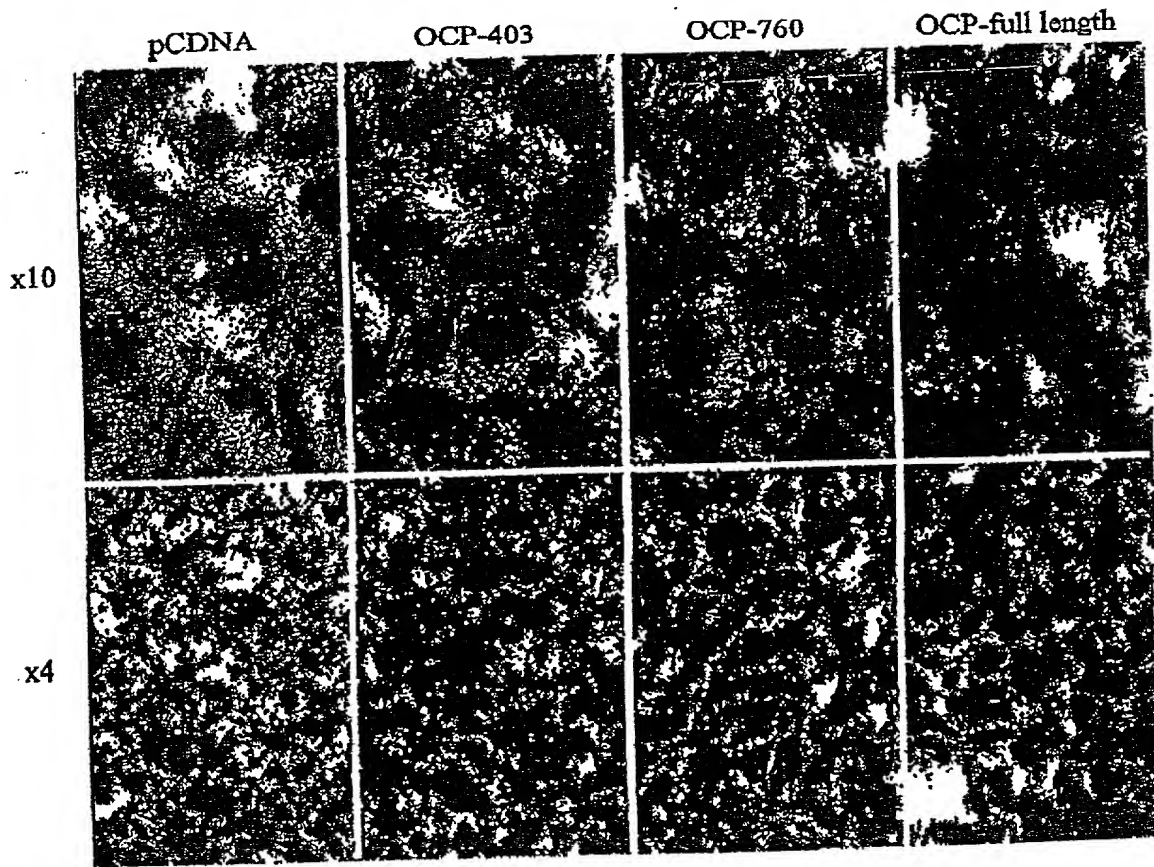


Figure 36

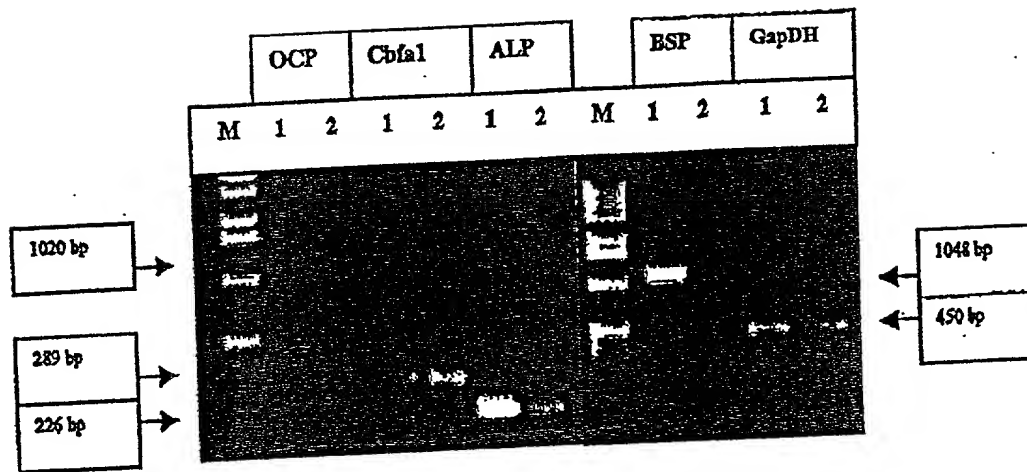


Figure 37

x 4

x 10

pCDNA
ROS stable line

OCP
ROS stable line



DATE	TIME	LOCATION	WIND	TEMP	WAVE	SEA	WIND	TEMP	WAVE	SEA
1955-05-15	08:00	1000	10	22	10	10	10	22	10	10
1955-05-15	12:00	1000	10	22	10	10	10	22	10	10
1955-05-15	16:00	1000	10	22	10	10	10	22	10	10
1955-05-15	20:00	1000	10	22	10	10	10	22	10	10
1955-05-16	04:00	1000	10	22	10	10	10	22	10	10
1955-05-16	08:00	1000	10	22	10	10	10	22	10	10
1955-05-16	12:00	1000	10	22	10	10	10	22	10	10
1955-05-16	16:00	1000	10	22	10	10	10	22	10	10
1955-05-16	20:00	1000	10	22	10	10	10	22	10	10
1955-05-17	04:00	1000	10	22	10	10	10	22	10	10
1955-05-17	08:00	1000	10	22	10	10	10	22	10	10
1955-05-17	12:00	1000	10	22	10	10	10	22	10	10
1955-05-17	16:00	1000	10	22	10	10	10	22	10	10
1955-05-17	20:00	1000	10	22	10	10	10	22	10	10

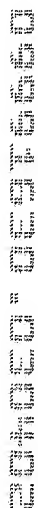
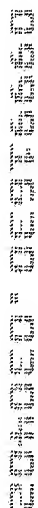
[illegible]

Figure 40

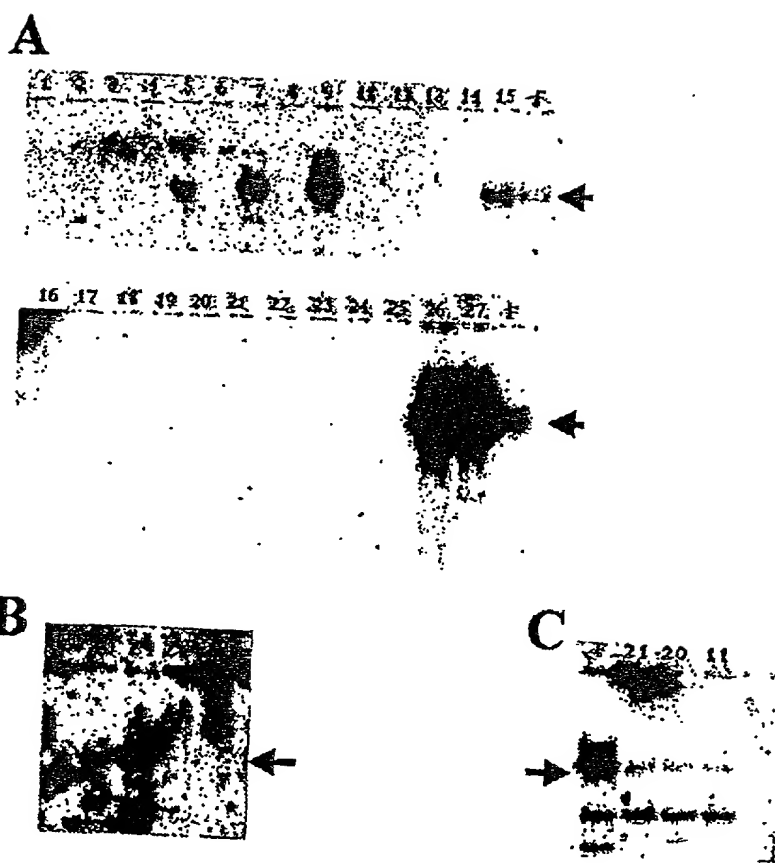


Figure 41

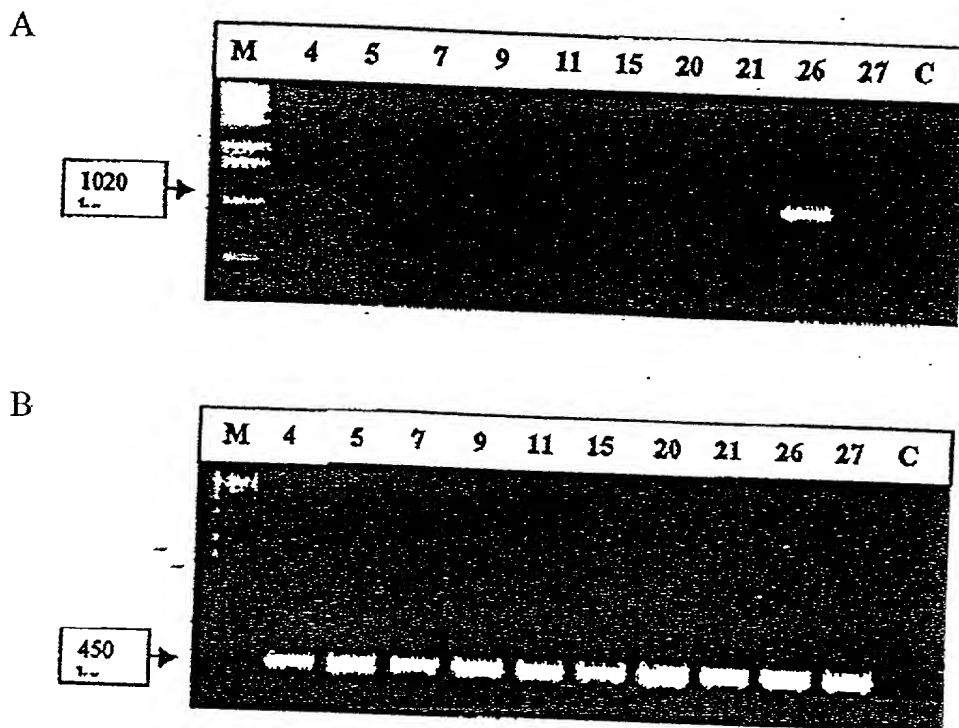
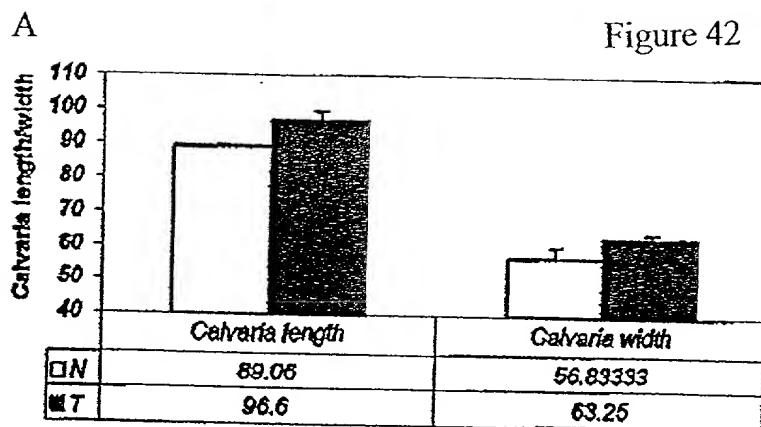


Figure 42



B

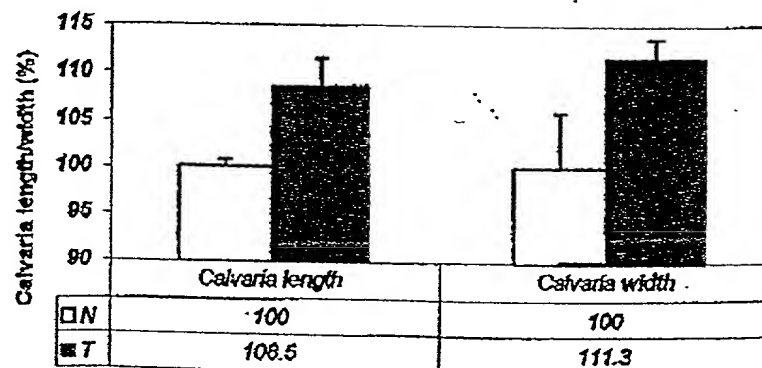


Figure 43

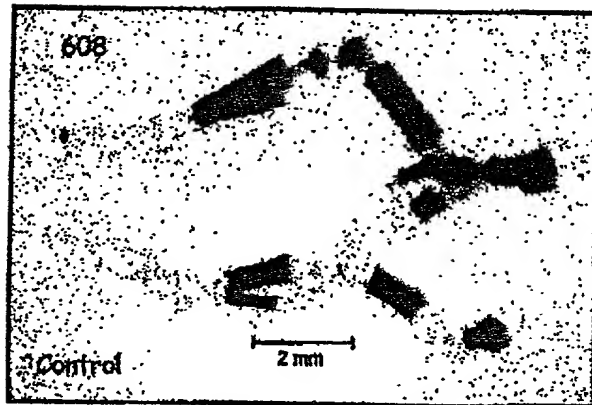


Figure 44

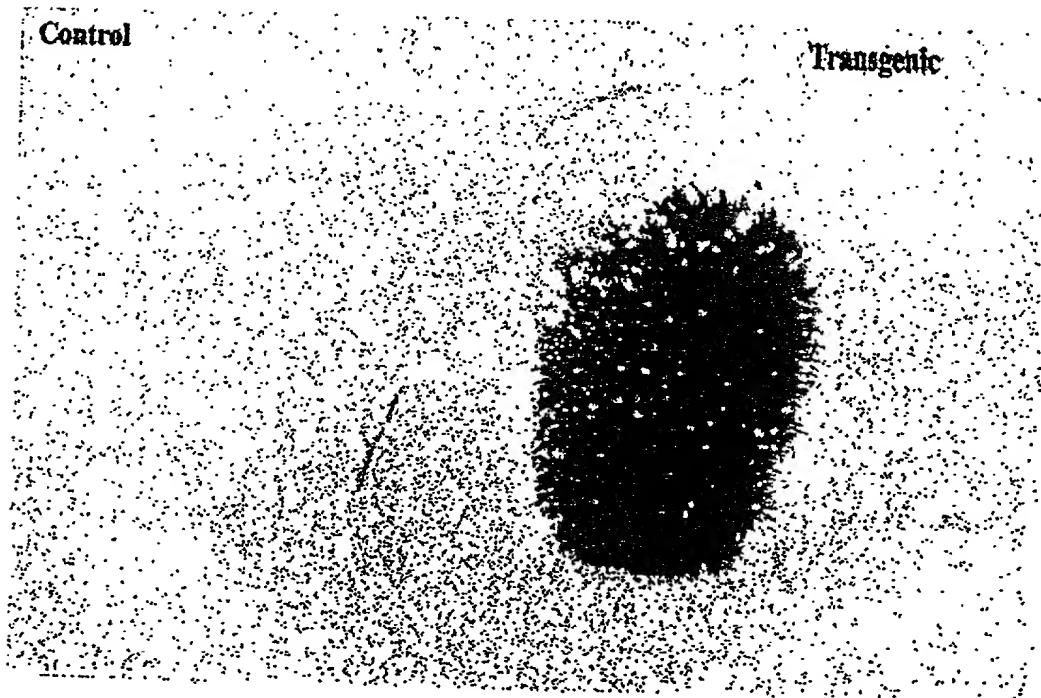
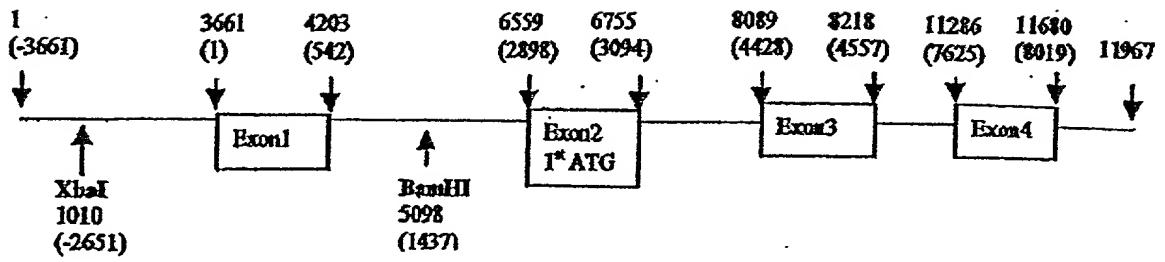


Figure 45



Clone 14C10 / ISE11

T7 orientation in pKS

Figure 46

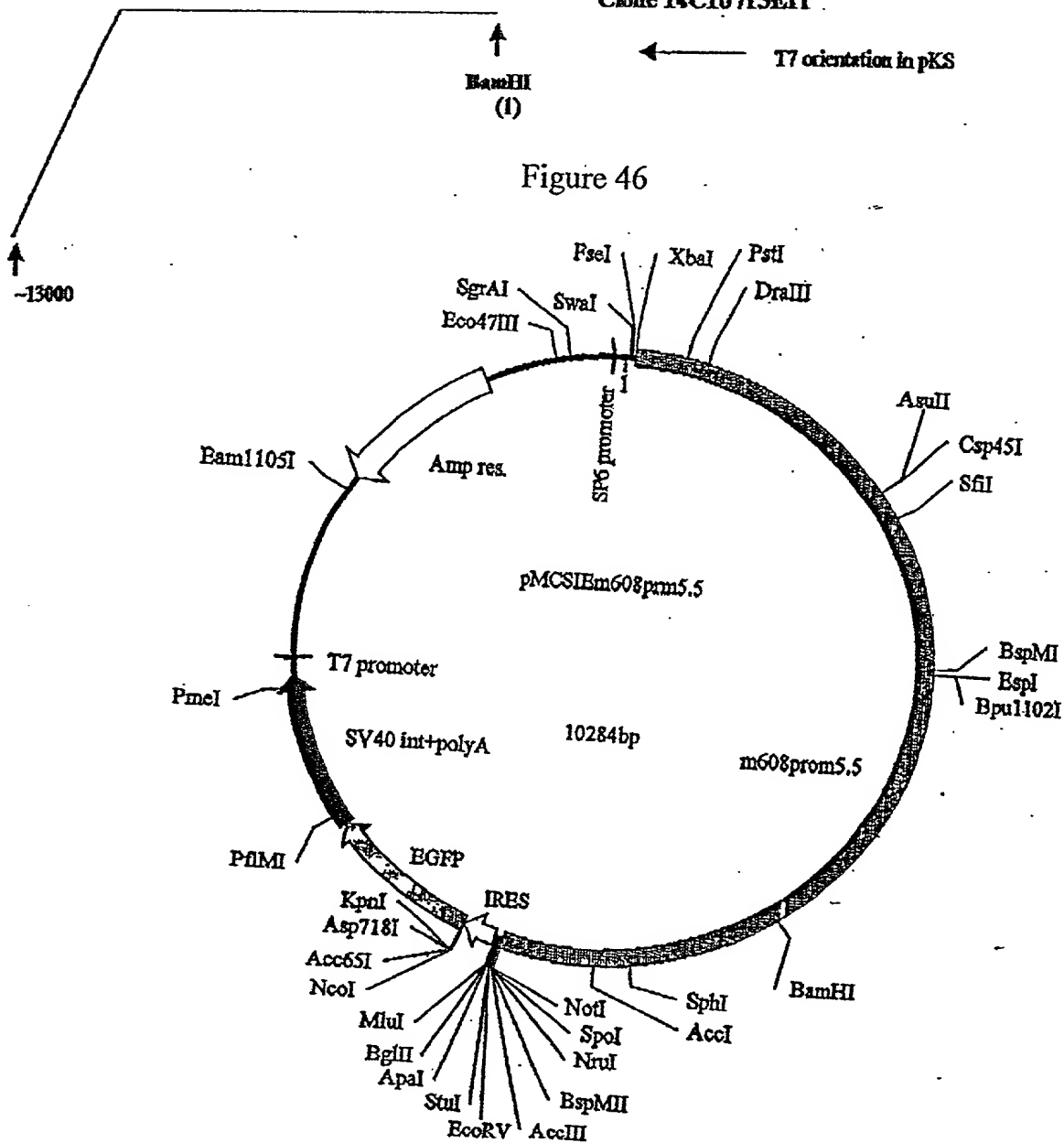


Figure 47

↓ (XbaI)

TCTAGAAGTAAATGATCCTGAGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGC
AATCATCTCTCTGTGGTTTGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA
ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT
GTAGCTGCCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCACATGCGGACAT
CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTOCAG
GGTTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTACCTTGATCATTAAATGACTAGAA
ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAGCCAGGCCAC
CTGAGCTTTGGCAGCTCCATTTCATTCTGTGGAGCCAGTCACGTGGGTTTGAATCAACTGT
ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTTCACAAGTCTTTTTA
GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAATAAACTGACTTTAAAAAATAA
AGCTTATGATACTTGTACAGAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTAT
TTTCATAATCCATCAATTAACAACTTTAGAAAATTTGTTAACACAAAGATCCCTAGGCC
CTGCCCTAGGATGGTCTGTATGGTGGGCTGAGAGATGGAGCTTAAGAACTTACTTGCTCC
AGGAGCACATCTTCAGAACATCTGCCTCAAAAACATTTATCCCAAATGCTCATCAAAGGCTC
ACTCACATGTGCTTCAACCACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGT
GGAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
ATAACCTCAACCACAACAAACAACGACAACAAACCCATAAAAAATTATCAGCGCAAAAT
GAACCTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCTTAAT
AGTCAAAATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATA
CTTGTAGATTAGAGGACACAAAGCAACAACCTACAGAGTGAAGTAATCCAGAGACTTAATG
TATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGATCTT
ACTAACACACACACTTACATAGAGAGATAAAGTGATAGATACGTTTGTITTTATCTTCATG
TAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAAGAGGAAGAGGAT
AGGTGCCTCTCCAGGGCGGAAGAGTACATTGAAAAGTCAGACACCATTGTGTAGATGTA
CCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCATATAGG
TGGAACAACATTATGAGCTAACCAAGTACCCCGAGCTCTTGACTCTAGCTGCATATATATC
AAAAGATGGCCTAATCGGCCATCACTGGAAGAGAGGCCATTGGACTTGCAAACTTTAT
ATGCCCCAGTACAGGGGAATACCAGGGCCAAAAAGGGGAGTGGGTGGGCAGGGGAGTG
GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATAC
CTAATAAAAAATGGAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCAC
TTCACTGCTATCTCAACTTGCAACAGAGGGGAGTCACAAACCCAGGACAAACACAGT
GATTGAAGCGTCTTTGAATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTAT
TGTGAAAACCTTACGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGC
ATATTTGGCACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACA
CACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGG
CATCATCTGCTCACTACCCACCCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
CTATTTTATGTCTAGGTTCCCTTCCCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA
AGAATTTGTAGCCCCGTAGCCAGCCCCGGGCCACAGAGCCTGCCACCAGACAGGAAAAGCC
CAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGGGTGGCGCTGGTGGT
CCCGGGTCGATGGCCCGCCATTCCCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCGG
CCCTCCTCCCGAGTGAAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTTCAGGGCGAGACAGAGGGTGGGATGGGCAGAAAG
GGTCCAGGAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAA
CATCGCTTCAGTGACTGAAGCCAGGCCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGA
CGCTTTCATAAACACTGATGACGTGTTTGTGCAAGCAAGCAATTTGAGGAGAAACGGCT
GGGACGTCGGGAAAGAAGGAGTGATCGATTAGTACTTGTAAAGTTTAGGTGAGTTGAGAAC
TAACCTAATCTATTGAGGGAGAAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCA
GCAATCAGATAAAGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAAATAACA
GAAGCGCCTCCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCCTTTTGTCCCGTTGGAAGCAAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGA

AGATTTCCTGCTCTGCTTATGTTACAAGAGGATTCAAAGCAAGACAGAAAGAGCTCAGG
TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATGTGGTA
TTCATTTACGGCTCTCTTCTTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
AOCGAGAGAGTATTCAAAAATATTTCAGTGATGTAACAGTGACAGTGTAGGCATAGAAAGTA
ATCATTAGTAAATCTTAATTTGGGTTAACTCATTCAACAGCTCCAGGTTGGGAGGGAT
CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
CTTGGCCATGCTCCCCATCACTGTGTCAAGTAAAGCAGAGGGGTGTTTCCAAGCAGAGAAAAG
AGCAGACAGTGTATGCCTGCAAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT
GTCTCCCGGTCAATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC
AGCATGTCTAGAAAGAAAATTGACCTTTTCTAGTGGGTTATTACAGCTGTAAAAGTAT
TTTGGAAAGGTTAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAATT
GATCATATAAAGAAAGTACATTCAATAATGCAATGTGAAAAATATATATAATTTTATCTAT
TACTGGTGCAAAGTTTTCTAAATTGCACATGTACTATTTTATATTTATAAAAAATATTTT
AAAATGTATATAAAAAGTGTAAGAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAAGCTATGAAGATATGTTTACAATAATAAGCCATTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
CTCTAGGGAATAAGAGTTCTCTTCTCCAGATTTCACTAGCATTTGTTGTCACTATTTATC
TTCTTGATGATGAGCATTATAAGTGGATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTCTTTCAGTTCACTAGTCTATTCAATTATTGGATAA
TTGGGGGATGGTGTAAATTTTTCAGTTCCTTATGGAATTCCAAAAACAAAAACAAAG
AAACAAACAAAAACCTCTGAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAAATTTATTGCTACATTGTTTATTATTACGACAGCCAAGAAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAGAGGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAGGAAAAACCCGTGTAGCTT
CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAAGCAAAAAACA
AGCTAGCTGGGGTGGGTAGGAGAGAAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
GGCACACGTTCAAACCAAGGGTGAAATCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
TTCGCCACTAACCAAGAAGCCATTGTCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTT
CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATAACAAAGTTGGCATGATGGATGG
TTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTTGGTGGTGTGTTGGTGGT
GCTGTGTTTTTCAATTTGTTCTTTTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
AGCTGGAACGATCTGGAAGAAGTTGGGGAAGAGAAAAATTTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAGGTTCAATTTGCCACAAAAAGGTGTGAATTAATTAAC
TAGTTACGACTCTTAAAGAAAAATATCCCAATTATCCAGAGTTGCTATGTATGCTGTGC
TAGGACTTTGCTTGAACCTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial

NotI site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCTCTTCTAACTCTCTCCCCATCTTGTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAAGTCTCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTCTGACCAATCCCCAAGTTATCACAAGGAGTTT
TTCTCAAACCTTTCTCTCCAGCAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCT
TCATTGGGAACCTGTGCTCCATCCAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCACTCTTTGTATCCCTCAAGAACACTGGGTCTGAAAAATTTAAACCCAAAG
AACTGTTTTTGTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTGTCTCCTCTTCAGTACCACCTGCTTTGCTGCTTTCTCAAAGAATCAATAAA
ACCAAAGTGATATAGATTTCATGGCAATTCCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGGAAGCAGTCAAGAGGCAGTTTCACAGAAACTTAGCCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCAGATTGTCCCTGAGATATTGTGGTG
GGTCTTGCACTGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAAATCTTCATGACAGTCAGGTATTGCAAACTACATCTGCATTA
GAAGAAAGTTGGTAAATCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAGGGGGGGGGTACT
GCTTTATTTCATGATAAGGTCTGCTAGAGCAGAAACCCCTGGGTGCTAGCTTTCAAGGTTTC
AAAGGTGTAGCATAAATTGTGACTAGAGTGTGAAATCTTTACCTGTCATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCTGTCAGTCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCTGCTGTGACACCTACTCCTTTGT
TTTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTGTGTCTGTGTCTTTTGTCAATTTCTGGCCAACTCCATGCCCACTGCCAGAACCCCTA
GTCTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTGTACACCCCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAAACCATCTCTTCTCGTTTGTCTAATATAGCAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCTGTATGTG
TGTCATTTGTTTCTTCAGTTCTGTGCCCCATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGGCTTTCAACGTTTTGCTG
TAAACAAGATATATGGTGGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC
CATTCTCATTCTGCTTCCACGCTGAGCATCTGTGTTTCCACAGGGACCACAAGAGCATTGA
AAAGTAGTGATTTATGACCTGCTTTGTCTGTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGTGACCTGTATCTGTGTTCTGGGTCTGCTCACTCATA
TTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCAACATTCTACTTCTGTAAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACCTCAAATATTTCAAATCTTTTAGCCCTCAGCCTGGGATTTCTTTG
ACTAAGTCTATTGATTTGGGAAGATCTCAGTGGTTAGGATTGCAGTTCATGATGTTTCATACG
TCAGGCTAAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTGCTGGGAATGTG
AGTGTTAGGGGGTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTCATAAGGAT
AGGTCTCATTCAACCACATGTTTATCATTTGGGATCTGCATTTTCTGATGCAAAATGATTT
ATTCTTCCAGAGCACTGGAATTGGGTTGAATCATCTTATAACGGCCAAAACATAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAAATGTGAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTGGCAGGATCTC
TTGGAACATCACACAAAAGAAATTCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTTCAAGCTTGACTTTGAG
GTGAAAGGAAAGCTCGCCAACTCTTTTATATGTAACACAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTGGGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCACTATAAAGCTTTATGAGAAATTT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTCTTCATTTACATTTT
AAATGCTATCCCAAAAGTCCCTATACTCCCCCTCACCTGCTCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCAATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCGAGCTGATGCAGTCAGATACAGACACCTAC
AGCCAAACAGTGAAGGAACCTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

204060 204060 204060

2049530-03403

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
AOCCTTGGGGCTCTCAGAGTCTGAACCAACCAACCATAGAACATTTCATGGGGCTGTACCCAG
GCCTCTCCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAAGTTCTGAGGTGTGCGGTGGAGGGTATACTCAGGGAGGCCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGGAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAAATAATTAAAAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAAAACATTGCATCCCCCCCCCAATCATTTCAGTGACTACGAATTAAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTAACTTCTTTGTGAECACGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTCTTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAAATAACAATGAAGAATAAAGACAACCTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCAOCTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTTATAAGAAAGTAAAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAACATAGAAAGTTGAAATTAAGAGAGTAGACAGGA
GAAGTAATAAGGCAAAATAATGAATATTTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCTACTTGAAAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGAACCAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

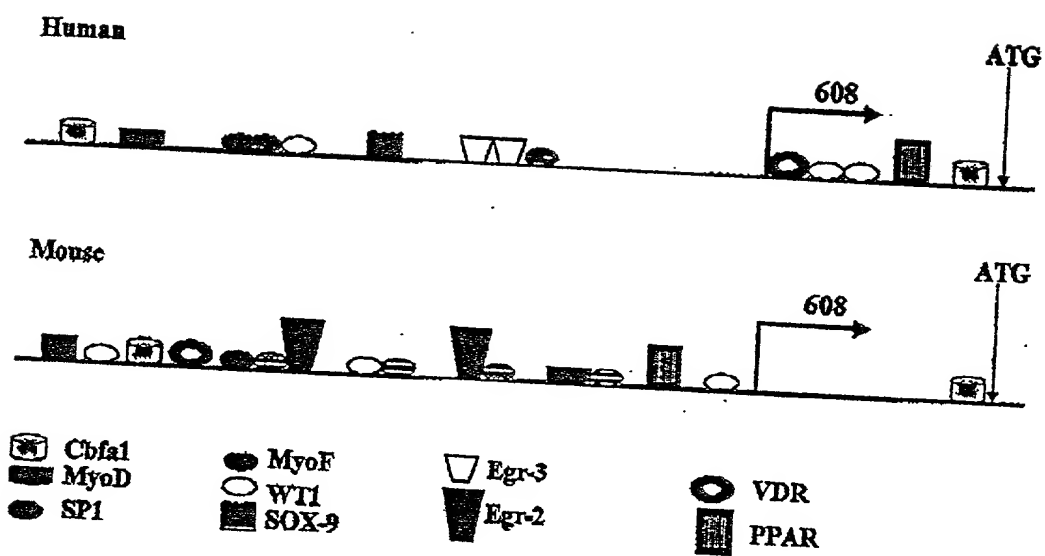


Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACCTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTTCATTTGCGGCTCTCTTCTTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGAACCAACCCAGATGCCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTCATATTA ACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTA ACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTA CTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCTGTCACACAGGTCTCCAGGG
T TAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAACTAGTA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCTGTGTGGAGCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTT
CACAAGTCTTTT TAGAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

20951650

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
TAAACTGACTTTAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAACCTTTAGAAA
TTTTGTAAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCCCTTAATAGTCAA
ATATGCAGAAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACACTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT
ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACCTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA
AAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACCTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCGCCCCCAACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCCCTTCCCTGCAATAGTACCTC
TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACCTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTACGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTGCGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTAAATCATCACGCAGAGAGTATTCAAAAATATTCAAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGCCATGCTCCCCATCACTGTGTCAGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGAAGGTTAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATTTTTTAAATG
TATATAAAAGTGTAAGAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCCTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

204060 0074650

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTCATTTT
GCCACAAAAAGGTGTGAATTAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCGAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTCTTAAATTAGTGATTTCATGGG
GCAGGGCCCAGTCCATTGTTCTGTGGTACCATTCTCAGGCCTATTAAAAAAAAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCTACCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAAGTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTCATTTTGGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAAGATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
TCAGAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAGGGAGGAAAGGAGGG
AGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGATGGGGAGGGAGGGAAGGAAG
GAAAGGGGGAGAAAGAAGAGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTTCACACACCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACATAAAGTACATAT
GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTACAGAGACTTTAAACATGAAACTGGAACCTATGTCTAGTGTTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTTCATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAATTT
TTTGCTTCTAATTCAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAATG
AGATGGTAGAAAGGAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAATCTAA
AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCAATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACCTATCTTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCAATTTAGAATGATATCCCAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALSVVLLILWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEEEQEQ
EEDGGSQLEKFPQLPQWSISLNMDEHGNMVNLVCDIKKPMDEVYKIHLNQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTGVRQAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYTQYSQTISTKDTRQARGRS
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPD SKFSIL
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVITGKNPGES
VTLP CNALAIPEAHL SWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
NQQGADHFTVGITVTKKGSGLPSKRGRRP GAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSL SLEVTPPFPAVSPPSASP
VQTVTSAEESSADVPLLGEHEHVLGTISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPT LISEPYEPSPTLHTLDTVYEKP THEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLET KSQPD EDKMKEDTFAHLTPPTI WVND S
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGMKEMSQTLQGGNM
LEGDPTHSRSSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETT VGTLLDKD TTTVTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGT PPRKHGKRPNKHRYPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDMTTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTD PPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSPS
TILMSLGQTTTTK PALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEG TQHMSGPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFT NKTL SF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLD FGPPAPPLLHTPQT TGSPSTN
LQNIPMVSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
TAETD TVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTD VRYG
GTLKLD CSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLV VKSVTDK

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted April 28, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVIFYGGDLKVDCVATGL
PNPEISWSLPDGLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMVRVRKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVQPPKING
NPNPITTVREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPAPIYYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNC
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNAA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPCPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRGTYYVCRMETEGPSVTSIPVIVIAYPPRITSEPTVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLTIQHAT
QRDAGFYKCMANKNILGSDSKTTYIHVF (SEQ ID NO: 21)

2020-09-23 10:44:40

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA
GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTT AGACTCTCCA
CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
TCCGTGGACC TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
AATCCAGAGG AATTCTGAAG TGTA AAAAGG ACAAAGCTTA TGAAGGCGGT
CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAAGTGC CCCAGTGGAG
CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
TCCAATGACC CGAGAAAACCT ATGAAAAGCT ATGGAAATTG ATAGCATACT
ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGTCTTTA
CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
GTGCTACTTT CTTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CCGTGACATT GCCTTGCAAT
GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

CGGAGCGCGCACTG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCTG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTCTCTGCTG TTTCTCCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACCTATACT TTTGCCTAGA
ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
 TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
 GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
 AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
 GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
 GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
 TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
 AGGAAGAAGC TGGTTCCTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
 GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
 CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCTA
 CTGCTGCCCC GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
 ATGTCTTTGG GACAAACCAC CACCCTAAG CCAGCACTTC CCAGTCCAAG
 AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
 TGGGGAATCC AGAAACAGAA GCAACCCAG TCAACAATGA AGGAACACAG
 CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
 ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
 GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
 GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
 AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
 TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
 TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
 ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
 CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTAATCCAAA
 GTGTTTGGA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
 TCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
 CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCAG
 ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
 TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
 CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
 CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
 CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
 GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
 TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
 TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
 AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
 GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
 ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
 CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
 CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTCCT
 GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
TGGAACTCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTTAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTT TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTTT GTATGCAAAG TTTTGTGTTGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGAAGCTTTT ACTGTATTTT
TAAGGTCAAT AACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGC
ATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGACAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
ACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCCTTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATC
TCTTTCATTAAGTTCTTATACCTTGTCTGATAACTTCCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCTTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGCTGCGATACAGGAGAAGCCAGATGTAATAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAG
CAGTGTCCACTTTGTCATGAACCTTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT
TTCCAGTGTGCCAAGCCAACCTTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT
AGTCTGCTTTTCATCTCTCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCACCCATTGCATTC
ACTGAAGAAAATGACTACATCGTCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAGATTAC
GGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGC
CACTTGCTTAGTGAAACACCCGACGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAACTACTTTA
CCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTG
GAACATACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCACCCACAC
GTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCAGTGAGGATGGACGGATCCTA
ATAGACAAAAGTGGAATAATTGGAACCTCCAGATGGCTGATAGTTTGGACACAGGCGTATATCACTGTATA
AGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTGAAGCC
TATCAGGAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAACACTTGATCTTCCATGCCATTCTACT
GGTATCCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGAC
AAGAAAGTTCTAAACAATGGCACATTAAGAATATTACAGGTACCCCGAAAGACCAAGGTATTATTCGC
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGGATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA
AGGCCCTTGGAGCATGATGGAGAAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT
AAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAAACACACCTCA
AGCACAAGTAAGAGGCACAACATATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGT
TTTAGGGAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCAGCTG
TTGGAGAAAAGCTAAAAAGAAATGCTATGCCAGACAAGCCGAGAAAATACCACAGTGAGCCACCCCAAGTG
GTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCCCTCAGGCATGCTCGCTCTACATGAGGAA
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA
TCTGATAGTCTCTATGACAAACATAAATTATGGCACAGAATTCTCTCTGTTGTGAATTCACAAATACTA
CCACCTGAAGAACCCACAGATTTCAAACCTGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA
AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTCTTTCCACTGCTA
CTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC
CCAATAACAGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAACTA
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC
AATCACTTCTATTCTCACACTACTCAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCT
GCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTTCAGACGCTTTGGGAGG
CAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGA
TACAGCATTTTCAGGTCAACAACCCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTG
CTCAATGTGACATGTCTGTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTT
CCAAGTGCTGCTCCCATCACCTTCCCAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACCT
CTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAAGAGAAAACACACCCACAAATAAAA
TATTTTCAGGACTGAAATTTCCCAAGTGAATCCAACTGGTGCAGTCATGACATATGCTCCAACATCCATA
CCCATGGAAAAAATTCACAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAGA
GATTCAAGTATTACATCGTCACTTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACA
AGGTTTTCAGAAAGGAAAATTCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAGGCAGATTA
AGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCTTTA
CCACAGAGACAAAGTTCCCTTTTCATTTTACCACACTTTTCAACAAGTGTGATGCAAAATTCATCTAAT
ACCTTGACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAG
CTTCCCTTCCACCCCTTAACCTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC
ATATCAACGCAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTCACTTATGAAACC
CAACAGAGAGATCTAGAGCACAACAATACAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGAC
CCAAACATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCA

20140607 15:56:00

TTCATCTATTCCCCACCAGAAAACACAACCTGGGATTTCAGCACAATCAGTTTTTCATTCAAGAACTCTT
AATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCT
GAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCATTAGACACTCCCATA
CCACCATTCTTGAGCAGCAGTGTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCA
GTTACTGACACACGTGGCGACTCCCATTTCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCC
TCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTA
CATTCTACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCTCCCTGG
GCAGAATACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGCCAAAAGCCAGAAGTAAGC
ATGTTGGCTACTACAGGCTGTCCGAGGCCACCACTCTTGTTCAGATTGGGATGGACAGAAGAACACA
AAGAAGAGTGACTTTTGATAAGAAACCAAGTTCAGAAAGCAACAACCTCCAACTCCCTTCCCTTTGACTCT
TTGTCTAGGTATATATTTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTACTATTCCAGCT
AACTCAGATGCCTTTCTTCCCTGTGAAGCTGTGTGAAATCCCCTGCCACCATTCAATGGACCAGAGTT
TCAGGACTTGATTTATCTAGAGGAAAACAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCC
ATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTGGCAGAGAC
CACCTTCATGTACCTTGTCTGTGGTTTCTATCCTCCCAGGATCCTGGAGAGACGTACCAAGAGATC
ACAGTTCATTCCGGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAAACCAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTT
GACGGAACATTGGTCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAAC
CCAGGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAA
AGGAGGCCAAGTCATTGTAGGCACCTGGGGTGAAGTTTAAACTGCCCTGTACTGCAAAAGGAACTCCT
CAGCCCAGCGTTTACTGGGTCTCTCTGATGGCAGTGAAGTGAACCAATTACAGTTTACCAATTCCAAG
TTGTTCTTATTTTCAAAATGGGACTTTGTATATAAGAAACCTAGCTCTTCAGACAGGGGCACCTTATGAA
TGCATTGCTACCAGTTCACCTGGTTCCGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACC
AGCCCCAGGATAGAAGCTGCATCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAAC
TGCTCAGCCACTGGGGAGCCCCAAACCCCAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCCGACCAG
TGGAGCTGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGT
GTCTACTTGTGTGTGGCAAGAAACAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTG
AAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTA
GATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAAC
AATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATTGGAATTTA
TACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGTCTATGCCCAGAACACCCTAGGGAAA
GATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAAC
AAGAGAATCAAAGCTGGAGACACAGCTGTCTTGACTGTGAGGTCACTGGGGATCCCAAAACCAAAATA
TTTTGGTTGCTGCCTTCCAATGACATGATTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACA
AACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCTGAAGGG
ACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCTCACAGCCCCATACTATGGAAGC
AGAATCACAGTCCATAAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTCAGATTACGCCGACTTT
ATCTGTGTGGCCCGAAATGAAGGTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTG
AGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTG
AATTGCTCTGTTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCC
AATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGG
GAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTA
GAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAACTCT
CTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTAT
GTAGTAGACAGGCCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAA
GCAACAGCTTATGACAGAGGAACTATATCTGTAAGGCTCAAATAGTGTGGTCATACACTGATTACT
GTTCCAGTAATGATTGTAGCTTACCCTCCCCGAATTACAAATCGTCCACCAGGAGTATTGTCACCAGG
ACAGGGGCAGCCTTTTCACTGCTGCTGCTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATG
CCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAA
GGTACCCTAGTCATTGAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCCA
CTGGTAGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 54

ORIGIN

1	MKVKGRGITC	LLVSFAVICL	VATPGGKACP	RRCACYMPTE	VHCTFRYLTS
51	IPDSIPPVNE	RINLGYNLSV	RLMETDFSGL	TKLELLMLHS	NGIHTIPDKT
101	FSDLQALQVL	KMSYNKVRKL	QKDTFYGLRS	LTRLHMDHNN	IEFINPEVFY
151	GLNFLRLVHL	EGNQLTKLHP	DTFVSLSYLQ	IFKISFIKFL	YLSDNFLTSL
201	PQEMVSYMPD	LDSLYLHGPN	WTCDCHLKWL	SDWIOQEPDV	IKCKKDRSPS
251	SAQQCPLCMN	PRTSKGKPLA	MVSAAAFQCA	KPTIDSSLKS	KSLTILEDSS
301	SAFISPQGM	APFGSLTLNM	TDQSGNEANM	VCSIQKPSRT	SPIAFTEEND
351	YIVLNTSFST	FLVCNIDYGH	IQPVWQILAL	YSDSPLILER	SHLLSETPQL
401	YYKYKQVAPK	PEDIFTNIEA	DLRADPSWLM	QDQISLQLNR	TATTFSTLQI
451	QYSSDAQITL	PRAEMRPVKH	KWTMISRDNN	TKLEHTVLVG	GTVGLNCPGQ
501	GDPTPHVDWL	LADGSKVRAP	YVSEDGRILI	DKSGKLELQM	ADSFDTGVYH
551	CISSNYDDAD	ILTYRITVVE	PLVEAYQENG	IHHTVFIGET	LDLPCHSTGI
601	PDASISWVIP	GNVLYQSSR	DKKVLNNGTL	RILQVTPKDQ	GYRCVAANP
651	SGVDFLIFQV	SVKMKGQRP	EHDGETEGSG	LDESNPPIAH	KEPPGAQLRT
701	SALMEAEVGK	HTSSTSQRHN	YRELTLQRRG	DSTHRRFREN	RRHFPPSARR
751	IDPQHWALL	EKAKKNAMPD	KRENTTVSPP	PVVTQLPNIP	GEEDDSSGML
801	ALHEEFMVPA	TKALNLPART	VTADSRTISD	SPMTNINYGT	EFSPVVSQI
851	LPPEEPTDFK	LSTAIKTTAM	SKNINPTMSS	QIQGTTNQHS	STVFPLLLGA
901	TEFQSDSQMG	RGREHFQSRP	PITVRTMIKD	VNVKMLSSTT	NKLLLESVNT
951	TNSHQTSVRE	VSEPRHNHFI	SHTTQILSTS	TFPSDPHTAA	HSQFPIPRNS
1001	TVNIPLFRRF	GRQRKIGGRG	RIISPYRTPV	LRRHRSIFR	STTRGSSEKS
1051	TTAFSATVLN	VTCLSCLPRE	RLTTATAALS	FPSAAPITFP	KADIARVPSE
1101	ESTTLVQNPL	LLLENKPSVE	KTTPTIKYFR	TEISQVTPTG	AVMTYAPTSI
1151	PMEKTHKVNA	SYPRVSSSTNE	AKRDSVITSS	LSGAITKPPM	TIIAITRFSR
1201	RKIPWQQNFV	NNHNPKGRLR	NQHKVSLQKS	TAVMLPKTSP	ALPQRQSSPF
1251	HFTTLSTSV	QIPSNLTSTA	HHTTTKTHNP	GSLPTKKELP	FPPLNPMPLS
1301	IISKDSSTKS	IISTQTAIPA	TTPTFPASVI	TYETQTERS	AQTIQREQEP
1351	QKKNRTPDNI	SPDQSSGFTT	PTAMTPPALA	FTHSPPENTT	GISSTISFHS
1401	RTLNLTDVIE	ELAQASTQTL	KSTIASSETTL	SSKSHQSTTT	RKASLDTPIP
1451	PFLSSSATLM	PVPISPPFTQ	RAVTDTRGDS	HFRMTNTTVV	KLHESSRHNL
1501	QMPSSQLEPL	TSSTSNLLHS	TPMPALTTVK	SQNSKLTPSP	WAQYQFWHKP
1551	YSDIAEKGGK	PEVSMLATG	LSEATTLVSD	WDGQKNTKKS	DFDKKPVQEA
1601	TTSKLLPFDS	LSRYIFEKPR	IVGGKAASFT	IPANSDAFLP	CEAVGNPLPT
1651	IHWTRVSGLD	LSRGNQNSRV	QVLPNGTSLI	QRVEIQDRGQ	YLCSASNLFQ
1701	TDHLHVTLVS	VSYPPIRLER	RTKEITVHSG	STVELKCRAE	GRPSPTVTWI
1751	LANQTVVSES	SQGSROAVVT	VDGYLVJLNL	SYNDRSFIYC	VASNPPGGQDS
1801	LLVKIQVIAA	PPVILEQRRQ	VIVGTWGESL	KLPCTAKGTP	QPSVYVWLSD
1851	GTEVKPLQFT	NSKLFLFSNG	TLYIRNLASS	DRGTYECIAT	SSTGSERRVV
1901	MLTMEERVTS	PRIEASQKR	TEVNFGDKLL	LNCSATGEPK	PQIMWRLPSK
1951	AVVDQWSWIH	VYPNGSLFIG	SVTEKDSGVY	LCVARNKMGD	DLILMHVSLR
2001	LKPAKIDHKQ	YFRKQVLHGK	DFQVDCASG	SPVPEISWSL	PDGTMNNAM
2051	QADDSGHRTR	RYTLFNNGTL	YFNKVGVAEE	GDYTCYAQNT	LGKDEMKVHL
2101	TVITAAPRIR	QSNKTNKRIK	AGDTAVLDCE	VTGDPKPKIF	WLLPSNDMIS
2151	FSIDRYTFHA	NGSLTINKVK	LLDSGEYVCV	ARNPSGDDTK	MYKLDVVSKF
2201	PLINGLYTNR	TVIKATAVRH	SKKHFDCAE	GTPSPEVMWI	MPDNIFLTAP
2251	YYGSRITVHK	NGTLEIRNVR	LSDSADFICV	ARNEGGSVL	VVQLEVLLEML
2301	RRPTFRNPFN	EKIVAQLGKS	TALNCSVDGN	PPPEIIWILP	NGTRFSNGPQ
2351	SYQYLIASNG	SFIISKTTRE	DAGKYRCAAR	NKVGYIEKLV	ILEIGQKPMI
2401	LTYAPGTVKG	ISGESLSLHC	VSDGIPKPNI	KWTMPSGYVV	DRPQINGKYI
2451	LHDNGTLVIK	EATAYDRGNY	ICKAQNVSCH	TLITVPMIV	AYPPRITNRP
2501	PRSIIVTRTGA	AFQLHCVAGL	VPKPEITWEM	PDHSLSLSTAS	KERTHGSEQL
2551	HLQGTLVION	PQTSDSGIYK	CTAKNPLGSD	YAATYIQVI*	

bioRxiv preprint doi: <https://doi.org/10.1101/2019.04.10.334400>; this version posted April 10, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible]

MQVRGREVSGLLISLTAVCLVVTGPSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDFTFVLSYLQ
IFKTSFIKYLFLSDNFLTSLPKEMVSYPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
SASTSPQDFIEPFGSLSNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
NLVCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWMTILMMNN
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKQGHYQCVAANPSGADFSSFKV
SVQ

Figure 56

pCM-H-608-663-N-term

